

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
24 February 2005 (24.02.2005)

PCT

(10) International Publication Number  
**WO 2005/017101 A2**

(51) International Patent Classification<sup>7</sup>: C12N

[GR/US]; 5915 Tamar Drive, Columbia, MD 21045 (US).  
CHIORINI, John, A. [US/US]; 2604 Loma St., Silver  
Spring, MD 20902 (US).

(21) International Application Number:  
PCT/US2004/015534

(22) International Filing Date: 18 May 2004 (18.05.2004)

(74) Agents: SPRATT, Gwendolyn, D. et al.; Needle & Rosen-  
berg, P.C., Suite 1000, 999 Peachtree Street, Atlanta, GA  
30309-3915 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/472,066 19 May 2003 (19.05.2003) US

(81) Designated States (*unless otherwise indicated, for every  
kind of national protection available*): AE, AG, AL, AM,  
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,  
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,  
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,  
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,  
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,  
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,  
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,  
ZW.

(71) Applicant (*for all designated States except US*): THE  
GOVERNMENT OF THE UNITED STATES OF  
AMERICA, as represented by THE SECRETARY, DE-  
PARTMENT OF HEALTH & HUMAN SERVICES,  
NATIONAL INSTITUTES OF HEALTH [US/US];  
Office of Technology Transfer, 6011 Executive Blvd, Suite  
325, Rockville, MD 20852 (US).

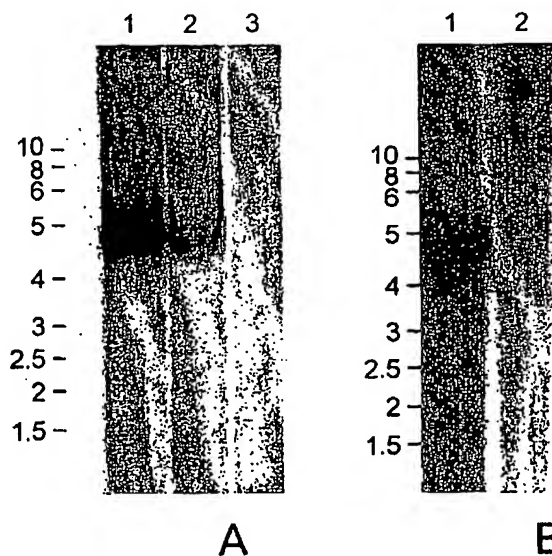
(84) Designated States (*unless otherwise indicated, for every  
kind of regional protection available*): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,  
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): BOSSIS, Ioannis

[Continued on next page]

(54) Title: AVIAN ADENOASSOCIATED VIRUS (AAAV) AND USES THEREOF



(57) Abstract: The present invention provides an Avian adeno-associated virus (AAAV) virus and vectors and particles derived therefrom. In addition, the present invention provides methods of delivering a nucleic acid to a cell using the AAAV vectors and particles. Methods of isolating the AAAV are provided.

WO 2005/017101 A2



FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

- *without international search report and to be republished upon receipt of that report*

## AVIAN ADENOASSOCIATED VIRUS (AAAV) AND USES THEREOF

### BACKGROUND OF THE INVENTION

5

#### Field of the Invention

The present invention provides avian adeno-associated virus (AAAV) and vectors derived therefrom. Thus, the present invention relates to AAAV vectors for and methods of delivering nucleic acids to cells of subjects.

10

#### Background Art

To date, eight AAV isolates (AAV<sub>1-8</sub>) have been, characterized and sequenced (2, 4, 19, 20, 25, 32, 51, 56) with AAV2 being the most extensively studied. AAV virions are approximately 20-25 nm in diameter and are composed of a mixture of assembled proteins (VPs) that encapsidate a linear ~4.7 kb single stranded DNA of plus or minus polarity (7, 43). The genome of AAVs is flanked by inverted terminal repeats (ITRs), which in the case of AAV2 are 145 nucleotides. The ITR is organized as three interrupted palindromes that can fold in an energetically favored T-shaped hairpin structure, which can exist in two orientations, termed flip and flop (42). The ITRs serve as origin of replication and contain *cis* acting elements required for rescue, integration, excision from cloning vectors and packaging (41, 42, 49 and 58).

20

The genetic map of the AAVs has been derived primarily from studies of AAV2 but is conserved in all serotypes (26, 27, 29, 36, 42, 45, 46, 58, 60, and 64). Two major open reading frames (*rep* and *cap* ORFs) and three transcriptional active promoters (*P<sub>5</sub>*, *P<sub>19</sub>*, *P<sub>40</sub>*) have been identified in the genome of AAV2. The *P<sub>5</sub>* and *P<sub>19</sub>* promoters encode for the nonstructural replication proteins Rep78 and Rep 68 and Rep 52 and Rep 40, respectively. Due to differential splicing, Rep78 and Rep52 have different C termini from Rep68 and Rep40. Transcription initiation from two promoters results in

25

30

Rep78 and Rep68 having different N termini from Rep52 and Rep 40. The *P<sub>40</sub>* promoter transcribes two alternatively spliced mRNAs. The major mRNA species encodes for the major capsid protein VP3 from a conventional AUG codon and the minor capsid protein VP2 from an upstream in frame ACG codon. The minor mRNA species encodes the entire *cap* ORF to produce the minor capsid protein VP1 (47). VP1, VP2 and VP3 are found in a ratio of 1:1:10, respectively, and this stoichiometry is generated by the high abundance of one of the mRNA species and the low translation efficiency from an ACG codon in the case of VP2 (14, 47, 55). Previous studies have indicated that VP2 and VP3 are sufficient for particle formation and accumulation of encapsidated ssDNA progeny, while VP1 is required for assembly of highly infectious particles (63, 64).

All four Rep proteins possess NTP binding activity, DNA helicase activity and nuclear localization sequences, however only Rep78/68 possess DNA binding ability (33, 34, 66). Mutant AAV defective for the synthesis of the small Rep proteins (Rep52/40) are able to replicate DNA but no ssDNA progeny is encapsidated (16). The ability of Rep78/68 to bind and nick DNA in a sequence and strand specific manner inside the ITR is essential in every phase of the AAV life cycle, namely DNA replication, AAV gene expression, rescue from the integrated state and self-excision from cloning vectors (29, 35, 44). Nicking of the DNA within the ITR at the terminal resolution site (*trs*) requires binding of Rep78/68 proteins to a motif composed of tandem repeats of GAGY.

Among AAV serotypes, AAV1, 4, 7 and 8 are believed to be of simian origin while AAV2, 3 and 5 are from humans. AAV6 was found in a human adenovirus preparation and is very similar to AAV1. AAVs have also been reported in other mammalian species including canines, bovine, ovine and equine (8). An avian AAV was first isolated from the Olson strain of quail bronchitis adenovirus (68). It was later found that 50% of adenoviral field isolates from chickens in US and Ireland contained AAVs serologically indistinguishable from the initial isolate (24). The AAV was



found to be 20 nm in diameter, serologically distinct from AAV<sub>1-4</sub>, did not agglutinate erythrocytes from several species tested and required adenovirus or herpes virus for replication (5, 68). In addition, AAV was found to inhibit replication of several avian adenovirus and herpes virus (5, 52, 53). Physicochemical studies revealed that the capsid of AAV consists of three VP proteins similar to other AAVs. The buoyant density of AAV in CsCl gradients (1.39-1.44 g/cm<sup>3</sup>) is similar to what have been reported for all AAVs (6, 30, 68).

The ability of AAV vectors to infect dividing and non-dividing cells and establish long-term transgene expression and the lack of pathogenicity has made them attractive for use in gene therapy applications. Recent evidence has indicated lack of cross competition in binding experiments suggesting that each AAV serotype may have a distinct mechanism of cell entry. Comparison of the *cap* ORFs from different serotypes has identified blocks of conserved and divergent sequence, with most of the later residing on the exterior of the virion, thus explaining the altered tissue tropism among serotypes (19-21, 48, 56). Vectors based on new AAV serotypes may have different host range and different immunological properties, thus allowing for most efficient transduction in certain cell types. In addition, characterization of new serotypes will aid in identifying viral elements required for altered tissue tropism.

20

Serological studies have provided evidence of avian adeno-associated virus infection in humans (69). Six percent of an unselected adult population was found positive for antibody to AAV by agar gel precipitation (AGP), and 15.6% was positive by virus neutralization (VN). Fourteen percent of poultry workers (industry or research) were positive for AAV antibody by AGP and 66% were positive by VN. In the same studies, no cross reaction was noted by AGP when antiserum to AAV was reacted against primate antigens of serotypes 1-4 or when antiserum to AAV serotypes 1-4 were reacted against AAV antigen. In addition, antiserum prepared against primate AAV1-4 did not neutralize the avian AAV. These results show that AAV is a distinct serotype and infections are not restricted to avian species but are found in the

30

human adult population.

Based on the genome organization and sequence homology among insect densovirus, rodent parvovirus and human dependovirus, it has been previously  
5 proposed these virus may have diverged from a common ancestor and evolved strictly in their hosts (3). However, the high sequence homology between avian autonomous parvovirus and primate AAVs and the epidemiological documentation of AAV transmission to humans provide evidence for host-independent evolution of at least some parvovirus genera. To better understand the relationship between the avian and  
10 the primate AAVs, the complete viral genome of AAV was cloned and sequenced and used to generate recombinant viral particles.

The present invention provides the first complete genomic AAV sequence. The genome of AAV is 4,694 nucleotides in length and has similar organization with  
15 that of other AAVs. The entire genome of AAV displays 56-65% identity at the nucleotide level with the other known AAVs. The AAV genome has inverted terminal repeats of 142 nucleotides with the first 122 forming the characteristic T-shaped palindromic structure. The putative Rep-binding element (RBE) consists of a tandem (GAGY)<sub>4</sub> repeat, and the putative terminal resolution site (*trs*), CCGGT/CG,  
20 contains a single nucleotide substitution relative to the AAV<sub>2</sub> *trs*. Surprisingly and in contrast to AAV5, the AAV ITR can be used as an origin of replication by either AAV5 or AAV2 Rep proteins for packaging. Thus the AAV ITR can act as a universal ITR. The Rep ORF of AAV displays 50-54 % identity at the amino acid level with the other AAVs, with most of the diversity clustered at the carboxyl and  
25 amino termini. Comparison of the capsid proteins of AAV and the primate dependoviruses indicate divergent regions are localized to surface exposed loops. Despite these sequence differences, recombinant AAV particles were produced carrying a *lacZ* reporter gene by co-transfection in 293T cells and transduction efficiency was examined in both chicken primary cells and several cell lines. This  
30 unique tropism allows AAV to be useful as a vector for the development of transgenic

animals and also allows for the vaccination of eggs as well as the preparation of recombinant proteins in avian cultures. The exposed regions of AAV are also sites for insertions of epitopes for the purpose of changing the tropism of the virus or antigen presentation. The present invention shows that AAV is the most divergent adeno-associated virus described to date, but maintains all the characteristics unique to the genera of dependovirus.

The present invention provides a vector comprising the AAV virus or a vector comprising subparts of the virus, as well as AAV viral particles. While AAV is similar to primate AAVs, the viruses are found herein to be physically and genetically distinct. These differences endow AAV with some unique properties and advantages which better suit it as a vector for gene therapy or gene transfer applications.

As shown herein, AAV capsid protein, again surprisingly, is distinct from primate capsid protein and exhibits different tissue tropism, thus making AAV capsid-containing particles suitable for transducing cell types for which primate AAVs are unsuited or less well-suited. AAV is serologically distinct and thus, in a gene therapy application, AAV would allow for transduction of a patient who already possesses neutralizing antibodies to primate isolates either as a result of natural immunological defense or from prior exposure to other vectors. AAV is also useful for gene transfer to other species for the development of transgenic animals or the production of vaccines and recombinant proteins in eggs. Thus, the present invention, by providing these new recombinant vectors and particles based on AAV, provides a new and highly useful series of vectors.

## SUMMARY OF THE INVENTION

The present invention provides a nucleic acid vector comprising a pair of avian adeno-associated virus (AAV) inverted terminal repeats and a promoter between the inverted terminal repeats.

The present invention further provides an AAV particle containing a vector comprising a pair of AAV inverted terminal repeats.

5 The present invention further provides an AAV particle containing a vector comprising a pair of AAV2 inverted terminal repeats.

The present invention further provides an AAV particle containing a vector comprising a pair of AAV5 inverted terminal repeats.

10 The present invention further provides an AAV1 particle containing a vector comprising a pair of AAV inverted terminal repeats.

The present invention further provides an AAV2 particle containing a vector comprising a pair of AAV inverted terminal repeats.

15 The present invention further provides an AAV3 particle containing a vector comprising a pair of AAV inverted terminal repeats.

20 The present invention further provides an AAV4 particle containing a vector comprising a pair of AAV inverted terminal repeats.

The present invention further provides an AAV5 particle containing a vector comprising a pair of AAV inverted terminal repeats.

25 The present invention further provides an AAV6 particle containing a vector comprising a pair of AAV inverted terminal repeats

The present invention further provides an AAV7 particle containing a vector comprising a pair of AAV inverted terminal repeats

30

The present invention further provides an AAV8 particle containing a vector comprising a pair of AAV inverted terminal repeats

The present invention further provides a dependovirus particle containing a  
5 vector comprising a pair of AAV inverted terminal repeats.

Additionally, the instant invention provides an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1 (AAAV genome). Furthermore, the present invention provides an isolated nucleic acid consisting essentially of the  
10 nucleotide sequence set forth in SEQ ID NO:1 (AAAV genome).

The present invention provides an isolated nucleic acid encoding an AAV Rep protein, for example, the nucleic acid as set forth in SEQ ID NO:2. Additionally provided is an isolated full-length AAV Rep protein or a unique fragment thereof.  
15 Additionally provided is an isolated AAV Rep 42 protein having the amino acid sequence set forth in SEQ ID NO:9, or a unique fragment thereof. Additionally provided is an isolated AAV Rep 52 protein having the amino acid sequence set forth in SEQ ID NO:5, or a unique fragment thereof. Additionally provided is an isolated AAV Rep 68 protein, having the amino acid sequence set forth in SEQ ID NO:7 or a  
20 unique fragment thereof. Additionally provided is an isolated AAV Rep 78 protein having the amino acid sequence set forth in SEQ ID NO:3, or a unique fragment thereof. The sequences for these proteins are provided below in the Sequence Listing and elsewhere in the application where the proteins are described.

25 The present invention further provides an isolated AAV capsid protein, VP1, having the amino acid sequence set forth in SEQ ID NO:11, or a unique fragment thereof. Additionally provided is an isolated AAV capsid protein, VP2, having the amino acid sequence set forth in SEQ ID NO:13, or a unique fragment thereof. Also provided is an isolated AAV capsid protein, VP3, having the amino acid sequence set  
30 forth in SEQ ID NO:15, or a unique fragment thereof.

The present invention additionally provides an isolated nucleic acid encoding AAV capsid protein, for example, the nucleic acid set forth in SEQ ID NO:10, or a unique fragment thereof.

5

The present invention further provides an AAV particle comprising a capsid protein consisting essentially of the amino acid sequence set forth in SEQ ID NO:11, or a unique fragment thereof.

10

Additionally, provided by the present invention is an isolated nucleic acid comprising an AAV p5 promoter having the nucleic acid sequence set forth in SEQ ID NO:22, or a unique fragment thereof.

15

The instant invention provides a method of screening a cell for infectivity by AAV comprising contacting the cell with AAV and detecting the presence of AAV in the cells.

20

The present invention further provides a method of delivering a nucleic acid to a cell comprising administering to the cell an AAV particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to the cell.

25

The present invention also provides a method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject.

30

The present invention also provides a method of delivering a nucleic acid to a cell in a subject comprising administering to the subject an AAV particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby

delivering the nucleic acid to a cell in the subject.

The instant invention further provides a method of delivering a nucleic acid to a cell in a subject having antibodies to other serotypes of AAV comprising administering  
5 to the subject an AAV particle comprising the nucleic acid, thereby delivering the nucleic acid to a cell in the subject.

The present invention also provides a method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV particle  
10 comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject.

The present invention also provides a method of delivering a nucleic acid to a cell in a subject comprising administering to the subject an AAV particle comprising  
15 the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to a cell in the subject.

The instant invention further provides a method of delivering a nucleic acid to a cell in a subject having antibodies to primate AAVs comprising administering to the  
20 subject an AAV particle comprising the nucleic acid, thereby delivering the nucleic acid to a cell in the subject.

#### BRIEF DESCRIPTION OF THE DRAWINGS

25 Figure 1 shows a Southern blot analysis of AAV nuclease resistant particles in 293T and LMH cells. A) 293T cells were transfected with pAAV alone (lane 3), pAAV plus pAd12 (lane 2) and pAAV plus infection with wt Ad (lane 1). B) LMH cells were transfected with pAAV alone (lane 2) or pAAV plus infection with FAV1 (lane 1). Viral DNA was isolated as described in Materials and Methods and  
30 fractionated on agarose gel before southern blot analysis with a <sup>32</sup>P-labeled pAAV

DNA.

Figure 2 shows the AAV ITR. The sequence of the ITR is shown in the hairpin conformation. The putative Rep binding site is boxed, while the putative *trs* is underlined and the cleavage site is indicated by an arrow.

5

Figure 3 is the sequence of an AAV genome. The genomes of AAV, AAV2, AAV4 and AAV5 were aligned using Clustal W. The sequences of the ITRs are presented in italics. The putative *trs* is indicated by vertical arrow and the putative RBS is underlined. Proposed transcription factor binding sites and the polyadenylation signal are also underlined. Proposed transcription initiation sites of the p5, p19 and p40 promoters and splice donor and acceptor sites are indicated by horizontal arrows. Initiation and termination codons are presented in bold letters.

Figures 4A and 4B illustrate comparisons of *rep* and *cap* ORFs. The *rep* and *cap* ORFs of AAV, AAV2, AAV4, AAV5 and Goose autonomous parvovirus (GP) were aligned using Clustal W. Identical amino acids are indicated by a dot. Dashes indicate gaps in the sequence added by the alignment program. A) Horizontal arrows indicate the initiator codon of the p5 and p19 Rep proteins. The Rep endonuclease site established by Tyr155 and the tetrahedrally coordinated Asp24, Glu83, His90 and His92 are presented in bold letters and are over lined by an asterisk. The region important for Rep multimerization, the ATP binding site and the basic amino acids of the nuclear localization signal are underlined. The zinc finger motifs in the carboxy terminus are underlined and the coordinating cystine and histidine residues are indicated by dots. B) The theoretical initiator codons of VP2 and VP3 are indicated in bold letters. Regions that have been proposed to be on the surface of AAV2 are underlined and divergent regions are boxed. The heparin binding region in the capsid of AAV2 is also indicated.

Figures 5A and 5B show vector constructs for generation of recombinant AAV virus and transduction of chicken fibroblasts. A) Wild type AAV, vector

30



plasmid (pA3Vbgal) and production yields of rAAAV using helper plasmids providing the rep gene under control of CMV, MMTV or the native P5 promoter. The helper plasmids pCA3VRC, pMA3VRC, pA3VRC were individually co-transfected with pA3Vbgal and an adenovirus helper plasmid in 293T cells and rAAAV was produced as described in Material and Methods. The number of rAAAV genomes produced in each group was determined by quantitative PCR and is expressed as DNase resistant particle/cell (DPN/cell). ITR: inverted terminal repeats from AAV, RSV: Rous Sarcoma virus long terminal repeat promoter, CMV: cytomegalovirus immediate early promoter, MMTV murine leukemia virus long terminal repeat promoter,  $\beta$ -Gal:  $\beta$ -galactosidase gene, SV40-polyA: polyadenylation signal from SV40. B) Relative transduction efficiency of primary chicken embryonic fibroblasts (CEF) and immortalized chicken embryonic fibroblasts (DF1) with equal particles of rAAAV expressing LacZ.

Figure 6 shows results with neuraminidase indicating that while AAV5 is sensitive to sialic acid treatment AAV is not.

Figure 7 shows that while AAV2 is sensitive to heparin competition, AAV is not.

Figure 8 shows the role of terminal lactose in AAV binding by incubating virus with different conjugates that had either terminal lactose or sialic acid. AAV5 is sensitive to competition with sialic acid conjugates but AAV is not. However AAV is competed by terminal lactose conjugates confirming ERCR lectin result.

Figures 9A and 9B show that treatment with tunicamycin blocks virus binding and transduction, suggesting that glycosylation is N-linked. AAV5 is the control.

Figure 10 confirms that glycoprotein is involved in AAV binding and transduction, cells were treated with broad specificity protease, trypsin. Like AAV5,

trypsin minimally effects virus binding. However treatment with low levels of tunicamycin dramatically increased the inhibition in binding observed with trypsin treatment..

5        Figures 11A and 11B show that the linkage is probably not an O-linkage.

Figures 12A and 12B show results with a series of N-linked inhibitors: NB-DNJ is a specific inhibitor of ER glucosidase I, II, and glycolipid; NB-DGJ glycolipid synthesis inhibiting properties as NB-DNJ; DNJ inhibits glucosidase 1,2; Fumonisin B1  
10        is an inhibitor of ceramide synthesis; and PDMP is an inhibitor of glycosphingolipid synthesis.

Figures 13A and 13B show results using several lectins and confirms previous results that sialic acid is not important (WGA vs WGA-s). *Erythrina corralodendron*  
15        (ERCL) which binds terminal poly lactose does inhibit suggesting that the virus is binding terminal lactose.

#### DETAILED DESCRIPTION OF THE INVENTION

20        As used in the specification and in the claims, "a" can mean one or more, depending upon the context in which it is used. The terms "having" and "comprising" are used interchangeably herein, and signify open ended meaning.

The present application provides a recombinant avian adeno-associated virus  
25        (AAAV). This virus has one or more of the characteristics described below. The compositions of the present invention do not include wild-type AAAV. The methods of the present invention can use either wild-type AAAV or recombinant AAAV-based delivery.

The present invention provides novel AAV particles, recombinant AAV vectors, recombinant AAV virions and novel AAV nucleic acids and polypeptides. An AAV particle is a viral particle comprising an AAV capsid protein. A recombinant AAV vector is a nucleic acid construct that comprises at least one  
5 unique nucleic acid of AAV. A recombinant AAV virion is a particle containing a recombinant AAV vector, wherein the particle can be either an AAV particle as described herein or a non-AAV particle. Alternatively, the recombinant AAV virion is an AAV particle containing a recombinant vector, wherein the vector can be either an AAV vector as described herein or a non-AAV vector. These vectors,  
10 particles, virions, nucleic acids and polypeptides are described below.

The present invention provides the nucleotide sequence of the AAV genome and vectors and particles derived therefrom. Specifically, the present invention provides a nucleic acid vector comprising a pair of AAV inverted terminal repeats  
15 (ITRs) and a promoter between the inverted terminal repeats. The rep proteins of AAV2 and AAV5 or AAV will bind to the AAV ITR and the AAV IRF can function as a universal origin or replication for packaging of recombinant AAV particles. The minimum sequence necessary for this activity is the TRS site (SEQ ID NO: 20) where Rep cleaves in order to replicate the virus. Minor modifications in an  
20 ITR are contemplated and are those that will not interfere with the hairpin structure formed by the ITR as described herein and known in the art. Furthermore, to be considered within the term e.g. it must retain the Rep binding site described herein. One of skill in the art would know how to modify an AAV ITR such that the hairpin structure is maintained and the Rep binding site is present. One of skill in the art could  
25 contemplate any ITR that contains a Rep binding site (SEQ ID NO: 21) and a trs site (SEQ ID NO: 20). Such an ITR could be utilized in any of the vectors described herein.

The D region of the AAV ITR, a single stranded region of the ITR, inboard of  
30 the TRS site, has been shown to bind a factor which depending on its phosphorylation

state correlates with the conversion of the AAV from a single stranded genome to a transcriptionally active form that allows for expression of the viral DNA. This region is conserved between AAV2, 3, 4, and 6 but is divergent in AAV5 and AAV. The D+ region (SEQ ID NO: 18) is the reverse complement of the D- region (SEQ ID NO: 19).

5

The promoter can be any desired promoter, selected by known considerations, such as the level of expression of a nucleic acid functionally linked to the promoter and the cell type in which the vector is to be used. That is, the promoter can be tissue/cell-specific. Promoters can be prokaryotic, eukaryotic, fungal, nuclear, mitochondrial,  
10 viral or plant promoters. Promoters can be exogenous or endogenous to the cell type being transduced by the vector. Promoters can include, for example, bacterial promoters, known strong promoters such as SV40 or the inducible metallothionein promoter, or an AAV promoter, such as an AAV p5 promoter. Additionally, chimeric regulatory promoters for targeted gene expression can be utilized. Examples of these  
15 regulatory systems, which are known in the art, include the tetracycline based regulatory system which utilizes the tet transactivator protein (tTA), a chimeric protein containing the VP16 activation domain fused to the tet repressor of *Escherichia coli*, the IPTG based regulatory system, the CID based regulatory system, and the Ecdysone based regulatory system (44). Other promoters include promoters derived from actin  
20 genes, immunoglobulin genes, cytomegalovirus (CMV), adenovirus, bovine papilloma virus, adenoviral promoters, such as the adenoviral major late promoter, an inducible heat shock promoter, respiratory syncytial virus, Rous sarcoma virus (RSV), etc.

Specifically, the promoter can be an AAV2 p5 promoter or an AAV5 p5  
25 promoter or an AAV p5 promoter. More specifically, the AAV p5 promoter can be at about the same location in SEQ ID NO: 1 as the AAV2 p5 promoter, in the corresponding AAV2 published sequence. Additionally, the p5 promoter may be enhanced by nucleotides 1-142 of SEQ ID NO:1. Furthermore, smaller fragments of the p5 promoter that retain promoter activity can readily be determined by standard  
30 procedures including, for example, constructing a series of deletions in the p5

promoter, linking the deletion to a reporter gene, and determining whether the reporter gene is expressed, *i.e.*, transcribed and/or translated. The promoter can be the promoter of any of the AAV serotypes, and can be the p19 promoter (SEQ ID NO: 23) or the p40 promoter set forth in the sequence listing as SEQ ID NO: 24.

5

It should be recognized that any errors in any of the nucleotide sequences disclosed herein can be corrected, for example, by using the hybridization procedure described below with various probes derived from the described sequences such that the coding sequence can be reisolated and resequenced. Rapid screening for point mutations can also be achieved with the use of polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP) (43). The corresponding amino acid sequence can then be corrected accordingly.

10

The AAV-derived vector of the invention can further comprise a heterologous nucleic acid functionally linked to the promoter. By "heterologous nucleic acid" is meant that any heterologous or exogenous nucleic acid, *i.e.* not normally found in wild-type AAV can be inserted into the vector for transfer into a cell, tissue or organism. By "functionally linked" is meant that the promoter can promote expression of the heterologous nucleic acid, as is known in the art, and can include the appropriate orientation of the promoter relative to the heterologous nucleic acid. Furthermore, the heterologous nucleic acid preferably has all appropriate sequences for expression of the nucleic acid. The nucleic acid can include, for example, expression control sequences, such as an enhancer, and necessary information processing sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites, and transcriptional terminator sequences.

15

20

25

The heterologous nucleic acid can encode beneficial proteins or polypeptides that replace missing or defective proteins required by the cell or subject into which the vector is transferred or can encode a cytotoxic polypeptide that can be directed, *e.g.*, to cancer cells or other cells whose death would be beneficial to the subject. The

30

heterologous nucleic acid can also encode antisense RNAs that can bind to, and thereby inactivate, mRNAs made by the subject that encode harmful proteins. The heterologous nucleic acid can also encode ribozymes that can effect the sequence-specific inhibition of gene expression by the cleavage of mRNAs. In one  
5 embodiment, antisense polynucleotides can be produced from a heterologous expression cassette in an AAV vector construct where the expression cassette contains a sequence that promotes cell-type specific expression (Wirak *et al.*, *EMBO* 10:289 (1991)). For general methods relating to antisense polynucleotides, see *Antisense RNA and DNA*, D. A. Melton, Ed., Cold Spring Harbor Laboratory, Cold  
10 Spring Harbor, NY (1988).

Examples of heterologous nucleic acids which can be administered to a cell or subject as part of the present AAV vector can include, but are not limited to the following: nucleic acids encoding secretory and non-secretory proteins, nucleic acids  
15 encoding therapeutic agents, such as tumor necrosis factors (TNF), such as TNF- $\alpha$ ; interferons, such as interferon- $\alpha$ , interferon- $\beta$ , and interferon- $\gamma$ ; interleukins, such as IL-1, IL-1 $\beta$ , and ILs -2 through -14; GM-CSF; adenosine deaminase; cellular growth factors, such as lymphokines; soluble CD4; Factor VIII; Factor IX; T-cell receptors; LDL receptor; ApoE; ApoC; alpha-1 antitrypsin; ornithine transcarbamylase (OTC);  
20 cystic fibrosis transmembrane receptor (CFTR); insulin; Fc receptors for antigen binding domains of antibodies, such as immunoglobulins; anti-HIV decoy tar elements; and antisense sequences which inhibit viral replication, such as antisense sequences which inhibit replication of hepatitis B or hepatitis non-A, non-B virus. The nucleic acid is chosen considering several factors, including the cell to be transfected. Where  
25 the target cell is a blood cell, for example, particularly useful nucleic acids to use are those which allow the blood cells to exert a therapeutic effect, such as a gene encoding a clotting factor for use in treatment of hemophilia. Another target cell is the lung airway cell, which can be used to administer nucleic acids, such as those coding for the cystic fibrosis transmembrane receptor, which could provide a gene therapeutic  
30 treatment for cystic fibrosis. Other target cells include muscle cells where useful

nucleic acids, such as those encoding cytokines and growth factors, can be transduced and the protein the nucleic acid encodes can be expressed and secreted to exert its effects on other cells, tissues and organs, such as the liver. Furthermore, the nucleic acid can encode more than one gene product, limited only, if the nucleic acid is to be packaged in a capsid, by the size of nucleic acid that can be packaged.

Furthermore, suitable nucleic acids can include those that, when transferred into a primary cell, such as a blood cell, cause the transferred cell to target a site in the body where that cell's presence would be beneficial. For example, blood cells such as TIL cells can be modified, such as by transfer into the cell of a Fab portion of a monoclonal antibody, to recognize a selected antigen. Another example would be to introduce a nucleic acid that would target a therapeutic blood cell to tumor cells. Nucleic acids useful in treating cancer cells include those encoding chemotactic factors which cause an inflammatory response at a specific site, thereby having a therapeutic effect.

Cells, particularly blood cells, muscle cells, airway epithelial cells, brain cells and endothelial cells having such nucleic acids transferred into them can be useful in a variety of diseases, syndromes and conditions. For example, suitable nucleic acids include nucleic acids encoding soluble CD4, used in the treatment of AIDS and  $\alpha$ -antitrypsin, used in the treatment of emphysema caused by  $\alpha$ -antitrypsin deficiency. Other diseases, syndromes and conditions in which such cells can be useful include, for example, adenosine deaminase deficiency, sickle cell deficiency, brain disorders such as Alzheimer's disease, thalassemia, hemophilia, diabetes, phenylketonuria, growth disorders and heart diseases, such as those caused by alterations in cholesterol metabolism, and defects of the immune system.

As another example, hepatocytes can be transfected with the present vectors having useful nucleic acids to treat liver disease. For example, a nucleic acid encoding OTC can be used to transfect hepatocytes (*ex vivo* and returned to the liver or *in vivo*) to treat congenital hyperammonemia, caused by an inherited deficiency in OTC.

Another example is to use a nucleic acid encoding LDL to target hepatocytes *ex vivo* or *in vivo* to treat inherited LDL receptor deficiency. Such transfected hepatocytes can also be used to treat acquired infectious diseases, such as diseases resulting from a viral infection. For example, transduced hepatocyte precursors can be used to treat viral hepatitis, such as hepatitis B and non-A, non-B hepatitis, for example by transducing the hepatocyte precursor with a nucleic acid encoding an antisense RNA that inhibits viral replication. Another example includes transferring a vector of the present invention having a nucleic acid encoding a protein, such as  $\alpha$ -interferon, which can confer resistance to the hepatitis virus.

For a procedure using transfected hepatocytes or hepatocyte precursors, hepatocyte precursors having a vector of the present invention transferred in can be grown in tissue culture, removed from the tissue culture vessel, and introduced to the body, such as by a surgical method. In this example, the tissue would be placed directly into the liver, or into the body cavity in proximity to the liver, as in a transplant or graft. Alternatively, the cells can simply be directly injected into the liver, into the portal circulatory system, or into the spleen, from which the cells can be transported to the liver via the circulatory system. Furthermore, the cells can be attached to a support, such as microcarrier beads, which can then be introduced, such as by injection, into the peritoneal cavity. Once the cells are in the liver, by whatever means, the cells can then express the nucleic acid and/or differentiate into mature hepatocytes which can express the nucleic acid.

The AAV-derived vector can include any normally occurring AAV sequences in addition to an ITR and promoter. Examples of vector constructs are provided below.

The present vector or AAV particle or recombinant AAV virion can utilize any unique fragment of these present AAV nucleic acids, including the AAV nucleic acids set forth in SEQ ID NOS: 1, 2, 4, 6, 8, 10, 12, 14, and 16-24. To be



unique, the fragment must be of sufficient size to distinguish it from other known sequences, most readily determined by comparing any nucleic acid fragment to the nucleotide sequences of nucleic acids in computer databases, such as GenBank. Such comparative searches are standard in the art. In particular, one of skill in the art will  
5 know how to distinguish an AAAV sequence from other AAV sequences. Therefore, the present invention provides AAAV nucleic acid sequences that are not found in other AAV sequences.

For example, one of skill in the art could perform alignments with an alignment  
10 program such as ClustalW or Blast2 where the parameters would be GAPOPEN or OPENGAP or OPEN GAP PENALTY : Penalty for the first residue in a gap (e.g., fasta defaults: -12 by with proteins, -16 for DNA). GAPEXT or EXTENDGAP or EXTEND GAP PENALTY : Penalty for additional residues in a gap (e.g. fasta defaults: -2 with proteins, -4 for DNA). Thus, it would be routine for one of skill in the art to utilize  
15 such alignment programs for identification of unique sequences as well as sequences that are 50%, 60%, 70%, 80%, 90%, 95% and 100% identical to the nucleic acid sequences described herein, as well as sequences that are 50%, 60%, 70%, 80%, 90%, 95% and 100% identical to the protein sequences described herein.

Typically, a unique fragment useful as a primer or probe will be at least about 8  
20 or 10, preferable at least 20 or 25 nucleotides in length, depending upon the specific nucleotide content of the sequence. Additionally, fragments can be, for example, at least about 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525,  
25 550, 575, 600, 625, 650, 675 or 700 nucleotides in length and can encode polypeptides or be probes. The nucleic acid can be single or double stranded, depending upon the purpose for which it is intended. Where desired, the nucleic acid can be RNA.

The present invention further provides an AAAV capsid protein to contain the  
30 vector. In particular, the present invention provides not only a polypeptide comprising

all three AAV coat proteins, *i.e.*, VP1, VP2 and VP3, but also a polypeptide comprising each AAV coat protein individually, SEQ ID NOS: 11, 13, and 15, respectively. Thus an AAV particle comprising an AAV capsid protein comprises at least one AAV coat protein VP1, VP2 or VP3. The present invention also provides particles comprising fragments of VP1, VP2 or VP3 that allow the particle to maintain AAV functionality and tropism. An AAV particle comprising an AAV capsid protein can be utilized to deliver a nucleic acid vector to a cell, tissue or subject. For example, the herein described AAV vectors can be encapsidated in an AAV5 capsid-derived particle and utilized in a gene delivery method. Furthermore, other viral nucleic acids can be encapsidated in the AAV particle and utilized in such delivery methods. For example, an AAV1-8 vector (e.g. a vector comprising an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7 or an AAV8 ITR and a nucleic acid of interest) can be encapsidated in an AAV particle and administered. Furthermore, an AAV chimeric capsid incorporating both AAV2 capsid and AAV capsid sequences can be generated, by standard cloning methods, selecting regions from the known sequences of each protein as desired. For example, particularly antigenic regions of the AAV capsid protein can be replaced with the corresponding region of the AAV2 capsid protein. In addition to chimeric capsids incorporating AAV2 capsid sequences, chimeric capsids incorporating AAV1, AAV3-8, and/or AAV5 capsid sequences can be generated, by standard cloning methods, selecting regions from the known sequences of each protein as desired. Alternatively a chimeric capsid can be made by the addition of a plasmid that expresses AAV1-8 capsid proteins at a ratio with the AAV capsid expression plasmid that allows only a few capsid proteins to be incorporated into the AAV particle. Thus, for example, a chimeric particle may be constructed that contains 6 AAV2 capsid proteins and 54 AAV capsid proteins if the complete capsid contains 60 capsid proteins.

The AAV capsid proteins can also be modified to alter their specific tropism by genetically modifying the capsid to comprise a specific ligand that binds to a cell surface receptor. Alternatively, the capsid can be chemically modified by conjugating a

ligand to a cell surface receptor. By genetically or chemically altering the capsids, the tropism can be modified to direct AAV to a particular cell or population of cells. The capsids can also be altered immunologically by conjugating the capsid to an antibody that recognizes a specific protein on the target cell or population of cells.

5

It has been recently reported that insertion of foreign epitopes (RGD motif, LH receptor targeting epitope) in certain regions of AAV2 capsid can redirect viral tropism. However, AAV2 naturally infects a wide variety of cell types and complete retargeting of rAAV2 would be difficult to achieve. For example, removal of the heparin binding activity, which is a major determinant of aav2 transduction in vitro, still results in AAV2 transduction of heart tissue in vivo. rAAV displays a more restrict tropism with preferential transduction of avian cells. Therefore AAV could be more easily engineered to specifically target certain cell types. Based on cryo-electron microscopy imaging of AAV2, 4 and 5, molecular modeling, and sequence alignments, we have identified regions in the capsid of AAV that are on the virus surface and could tolerate substitution. Two of these regions are aa 269-278 (PSGGDNNKF), and for some uses, more preferably aa 267-274 (QGPSGGD). There is evidence that the variable loop comprises QGPSGGD and that NNNKF may be conserved and may be important in structure. A substitution into PSGGDNNKF it did not assemble well and was not infectious in any cell. However, this type of insertion is useful for antigen presentation but not retargeting of the vector. Antigens presented in order arrays on the surface of viruses tend to be more antigenic than if they are presented in random. Another region is aa 454-463 (VSQAGSSGRA).

25 For insertion between aa 146-147

AAAV 146-LH nc: CGTCTTTGAGTCTTCCACCAGACCAAAG

AAAV 146-LH c:

30 CACTGCAGCACCTGCTACTACCACAAGAGCGCTCCGACCGGAGACAAG  
CG

For substitution at aa 267-274,

LHR-267-274F

5'CAACCACCTGTACAAACGAATCCACTGCAGCACCTGCTACTACCACAAGA  
GCAACAACAACAAATTCTTTGGATTC-3'

5 LHR-267-274R

5'GAATCCAAAGAATTTGTTGTTGTTGCTCTTGTGGTAGTAGCAGGTGCTGCA  
GTGGATTCGTTTGTACAGGTGGTTG-3'

For substitution at aa 269-278,

10 LHR-269-278F

CAAACGAATCCAAGGACACTGCAGCACCTGCTACTACCACAAGAGCTTTGG  
ATTCAGCACC

LHR-269-278R

GGTGCTGAATCCAAAGCTCTTGTGGTAGTAGCAGGTGCTGCAGTGTCTTG

15 GATTCGTTTG

For substitution at aa 454-463,

LHR-454-463F

TACCTCTGGGCTTTCAGCTCCCACTGCAGCACCTGCTACTACCACAAGAGCC

20 TTCATTACTCGCGGGCGAC

LHR-454-463R

GTCGCCCGCGAGTAATGAAGGCTCTTGTGGTAGTAGCAGGTGCTGCAGTGG  
GAGCTGAAAGCCCAGAGGTA

25 Other regions of the AAV capsid could also accommodate the substitution of amino acids that would allow for epitope presentation on the surface of the virus. All of these regions would have the following characteristics in common: 1) surface exposure, 2) ability to support a substitution of sequence to insert the epitope, 3) allows for capsid assembly. Examples of other insertion or substitution regions on the virus surface are  
30 the regions around T385-R394, S588-R601, T589-R600, S455-R462, S455-R463, T546-Q559, R550-T556, V329-I338, G708-T720, Ser 710-Y728 of VP1.

Because of the symmetry of the AAV particles, a substitution in one subunit of the capsid will appear multiple times on the capsid surface. For example the capsid is made of approximately 55 VP3 proteins (i.e., 50 VP3 is 90% of the capsid and there are 60 faces on an icosahedron). Therefore an epitope incorporated in the VP3 protein could be expressed 55 times on the surface of each particle increasing the likelihood of the epitope forming a stable interaction with its target. An epitope inserted upstream of the VP3 ORF may be presented in both the VP2 and VP1 proteins, or up to 10 times on the surface of each particle. In some cases this ligand density may be too high for functional binding or this high a density of epitope may interfere with capsid formation. The epitope density could be lowered by introducing another plasmid into the packaging system for production of recombinant particles and the ratio between the packaging plasmid with the modified VP protein and the wild type VP protein altered to balance the epitope density on the virus surface. Thus one example would be on an epitope that is targeted for the mounds at the 3 fold axis of symmetry. By mixing in 2 wild type rep and cap expression plasmids with 1 mutant rep and cap plasmid, instead of the mutant epitope being found in all three mounds at each three fold axis (60 times), it will only be present in 1 mound (20 times).

Epitopes can be incorporated into the virus capsid for the purpose of 1) altering the tropism of the virus, 2) blocking an immune response directed at the virus, 3) developing a host immune response to the epitope for the purpose of vaccination, and 4) catalyzing a reaction.

Examples of epitopes that can be added to AAV capsids include but are not limited to the following proteins and protein fragments:

LH receptor binding epitope: Photoaffinity labeling of the lutropin receptor with synthetic peptide for carboxyl terminus of the human choriogonadotropin alpha subunit. Kundu GC, Ji I, McCormick DJ, Ji TH. J Biol Chem. 1996 May

10;271(19):11063-6 (incorporated herein by reference);

RGD integrin binding epitope: Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. Pierschbacher MD, Ruoslahti

5 E. Nature. 1984 May 3-9;309(5963):30-3 (incorporated herein by reference);

CD13 binding epitope NGRAHA: Incorporation of tumor-targeting peptides into recombinant adeno-associated virus capsids. Grifman M, Trepel M, Speece P, Gilbert LB, Arap W, Pasqualini R, Weitzman MD. Mol Ther. 2001 Jun;3(6):964-75

10 (incorporated herein by reference) and F. Curnis, A. Sacchi, L. Borgna, F. Magni, A. Gasparri and A. Corti, Enhancement of tumor necrosis factor alpha antitumor immunotherapeutic properties by targeted delivery to aminopeptidase N (CD13). Nat. Biotechnol. 18 (2000), pp. 1185\*1190 (incorporated herein by reference);

15 Single chain antibody fragments: Q. Yang, Development of novel cell surface CD34-targeted recombinant adenoassociated virus vectors for gene therapy. Hum. Gene Ther. 9 (1998), pp. 1929\*1937 (incorporated herein by reference);

Endothelial cell binding epitope SIGYPLP: R. Pasqualini and E. Ruoslahti, 20 Organ targeting in vivo using phage display peptide libraries. Nature 380 (1996), pp. 364\*366 (incorporated herein by reference) and D. Rajotte, W. Arap, M. Hagedorn, E. Koivunen, R. Pasqualini and E. Ruoslahti, Molecular heterogeneity of the vascular endothelium revealed by in vivo phage display. J. Clin. Invest. 102 (1998), pp. 430\*437.(incorporated herein by reference);

25

Lung targeting peptide CGFECVRQCPERC: D. Rajotte and E. Ruoslahti, Membrane dipeptidase is the receptor for a lung-targeting peptide identified by in vivo phage display. J. Biol. Chem. 274 (1999), pp. 11593\*11598 (incorporated herein by reference);

30

Muscle targeting peptide ASSLNIA: T. I. Samoylova and B. F. Smith, Elucidation of muscle-binding peptides by phage display screening. Muscle Nerve 22 (1999), pp. 460\*466 (incorporated herein by reference);

- 5 Tumor endothelium targeting: W. Arap, R. Pasqualini and E. Ruoslahti, Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. Science 279 (1998), pp. 377\*380 (incorporated herein by reference);

Major immunogenic epitope for parvovirus B19 NISLDNPLENPSSFLDLVARIK:

- 10 K. Yoshimoto, A second neutralizing epitope of B19 parvovirus implicates the spike region in the immune response. J. Virol. 65 (1991), pp. 7056\*7060 (incorporated herein by reference);

- Serpin receptor ligand (KFNKPFVFLI): A small, synthetic peptide for gene delivery  
15 via the serpin-enzyme complex receptor. Patel S, Zhang X, Collins L, Fabre JW. J Gene Med. 2001 May-Jun;3(3):271-9 (incorporated herein by reference);

- Hemagglutinin (HA) 91-108: A retro-inverso peptide analogue of influenza virus hemagglutinin B-cell epitope 91-108 induces a strong mucosal and  
20 systemic immune response and confers protection in mice after intranasal immunization. Ben-Yedidia T, Beignon AS, Partidos CD, Muller S, Arnon R. Mol Immunol. 2002 Oct;39(5-6):323-31 (incorporated herein by reference);

- NDV epitope 447 to 455: Newcastle disease virus (NDV) marker vaccine: an  
25 immunodominant epitope on the nucleoprotein gene of NDV can be deleted or replaced by a foreign epitope. Mebatsion T, Koolen MJ, de Vaan LT, de Haas N, Braber M, Romer-Oberdorfer A, van den Elzen P, van der Marel P. J Virol. 2002 Oct;76(20):10138-46 (incorporated herein by reference);

- 30 RETANEF HIV-1 epitope vaccine candidate: A novel chimeric Rev, Tat, and Nef

(Retanef) antigen as a component of an SIV/HIV vaccine. Hel Z, Johnson JM, Tryniszewska E, Tsai WP, Harrod R, Fullen J, Tartaglia J, Franchini G. Vaccine. 2002 Aug 19;20(25-26):3171-86 (incorporated herein by reference); and

- 5 Catalytic single chain antibodies: Schultz, P.G. and Lerner, R.A., From molecular diversity to catalysis: lessons from the immune system 1995. Science 269, pp. 1835-1842 (incorporated herein by reference for its teaching of catalytic single chain antibodies responsible for the cleavage of protein substrates).

- 10 The capsids can also be assembled into empty particles by expression in mammalian, bacterial, fungal or insect cells. For example, AAV2 particles are known to be made from VP3 and VP2 capsid proteins in baculovirus. The same basic protocol can produce an empty AAV particle comprising an AAV capsid protein and also full particles.

15

- The herein described recombinant AAV nucleic acid derived vector can be encapsidated in an AAV particle. In particular, it can be encapsidated in an AAV1 particle, an AAV2 particle, an AAV3 particle, an AAV4 particle, an AAV5 particle, an AAV6 particle, and AAV7 particle or an AAV8 particle. A portion of any of the capsids, or a chimeric capsid particle as described above can be utilized, by standard methods using the appropriate capsid proteins in the encapsidation process, as long as the nucleic acid vector fits within the size limitation of the particle utilized. The encapsidation process itself is standard in the art. The AAV replication machinery, i.e. the rep initiator proteins and other functions required for replication, can be utilized to produce the AAV genome that can be packaged in an AAV1-8 capsid.
- 20
- 25

- The recombinant AAV virion containing a vector can also be produced by recombinant methods utilizing multiple plasmids. In one example, the AAV rep nucleic acid would be cloned into one plasmid, the AAV ITR nucleic acid would be cloned into another plasmid and the nucleic acid encoding a capsid (for example, an
- 30



AAV capsid from AAV1-AAV8) would be cloned on another plasmid. These plasmids would then be introduced into cells. The cells that were efficiently transduced by all three plasmids, would exhibit specific integration as well as the ability to produce AAV recombinant virus. Additionally, two plasmids could be used where the AAV rep nucleic acid would be cloned into one plasmid and the AAV ITR and AAV capsid would be cloned into another plasmid. These plasmids would then be introduced into cells. The cells that were efficiently transduced by both plasmids, would exhibit specific integration as well as the ability to produce AAV recombinant virus.

10

The capsid proteins of the present invention can have about 70% homology, about 75% homology, 80% homology, 85% homology, 90% homology, 95% homology, 98% homology, 99% homology, or even 100% homology to the protein having the amino acid sequence encoded by the nucleotides set forth in SEQ ID NOS:10, 12 or 14. The percent homology used to identify proteins herein, can be based on a nucleotide-by-nucleotide comparison or more preferable is based on a computerized algorithm as described herein. Variations in the amino acid sequence of the AAV capsid protein are contemplated herein, as long as the resulting particle comprising an AAV capsid protein remains antigenically or immunologically distinct from AAV1-8 capsid, as can be routinely determined by standard methods. Specifically, for example, ELISA and Western blots can be used to determine whether a viral particle is antigenically or immunologically distinct from AAV2 or the other serotypes. Furthermore, the AAV particle preferably retains tissue tropism distinction from AAV2, such as that exemplified in the examples herein. An AAV chimeric particle comprising at least one AAV coat protein may have a different tissue tropism from that of an AAV particle consisting only of AAV coat proteins, but is still distinct from the tropism of an AAV2 particle.

The invention further provides a recombinant AAV virion, comprising an AAV particle containing, *i.e.*, encapsidating, a vector comprising a pair of AAV

30

inverted terminal repeats. The recombinant vector can further comprise an AAV Rep-encoding nucleic acid. The vector encapsidated in the particle can further comprise an exogenous nucleic acid inserted between the inverted terminal repeats.

5           The invention further contemplates chimeric recombinant ITRs that contain a rep binding site and a TRS site recognized by that Rep protein. By "Rep protein" is meant all four of the Rep proteins, Rep 40, Rep 78, Rep 52, Rep 68. Alternatively, "Rep protein" could be one or more of the Rep proteins described herein. One example of a chimeric ITR would consist of an AAV D region (SEQ ID NOs: 18, 19 ), an  
10 AAV TRS site (SEQ ID NO: 20), an AAV2 hairpin and an AAV2 binding site. Another example would be an AAV D region, an AAV TRS site, an AAV3 hairpin and an AAV3 binding site. In these chimeric ITRs, the D region can be from AAV1-8. The hairpin can be derived from AAV 1-8. The binding site can be derived from any of AAV1-8. The D region and the TRS can be from the same serotype.

15           The chimeric ITRs can be combined with AAV Rep protein and any of the AAV serotype capsids to obtain recombinant virion. For example, recombinant virion can be produced by an AAV D region, an AAV TRS site, an AAV2 hairpin, an AAV2 binding site, AAV Rep protein and AAV1 capsid. This recombinant virion  
20 would possess the cellular tropism conferred by the AAV1 capsid protein and would possess the efficient replication conferred by the AAV Rep.

Other examples of the ITR, Rep protein and Capsids that will produce recombinant virus are provided in the list below but not limited to :

25           A<sub>ITR</sub> + A<sub>Rep</sub> + A<sub>Cap</sub>=virus  
A<sub>ITR</sub> + 5<sub>Rep</sub> + 1<sub>Cap</sub>=virus  
A<sub>ITR</sub> + 2<sub>Rep</sub> + 2<sub>Cap</sub>=virus  
A<sub>ITR</sub> + 7<sub>Rep</sub> + 3<sub>Cap</sub>=virus  
30 A<sub>ITR</sub> + 5<sub>Rep</sub> + 4<sub>Cap</sub>=virus

AITR + 5Rep + 6Cap=virus

1ITR + 1Rep + ACap=virus

2ITR + 2Rep + ACap=virus

4ITR + 4Rep + ACap=virus

5 5ITR + 5Rep + ACap=virus

6ITR + 6Rep + ACap=virus

(A= Avian, 1= AAV1, 2=AAV2, 3=AAV3, 4=AAV4, 5=AAV5, 6= AAV6)

10 In any of the constructs described herein, a promoter can be included. As used in the constructs herein, unless otherwise specified, Cap (capsid) refers to any of AAV VP1, AAV VP2, AAV VP3, combinations thereof, functional fragments of any of VP1, VP2 or VP3, or chimeric capsids as described herein. The ITRs of the constructs described herein, can be chimeric recombinant ITRs as described elsewhere  
15 in the application.

Conjugates of recombinant or wild-type AAV virions and nucleic acids or proteins can be used to deliver those molecules to a cell. For example, the purified AAV can be used as a vehicle for delivering DNA bound to the exterior of the virus.  
20 Examples of this are to conjugate the DNA to the virion by a bridge using poly-L-lysine or another charged molecule. Also contemplated are virosomes that contain AAV structural proteins (AAV capsid proteins), lipids such as DOTAP, and nucleic acids that are complexed via charge interaction to introduce DNA into cells.

25 Also contemplated by this invention is a method of delivering a DNA vaccine to a cell, comprising: administering a liposome comprising DNA conjugated to an AAV virion to a cell thus delivering the DNA vaccine to the cell.

High levels of humoral and cell-mediated immunity can be achieved via  
30 administration of DNA vaccines. Numerous studies have shown that immunization of

experimental animals with plasmid DNA encoding antigens from a wide spectrum of bacteria, viruses, protozoa and cancers leads to protective humoral and cell-mediated immunity (*Gregoriadis G.* "Genetic vaccines: strategies for optimization" *Pharm Res.* 15:661-70 (1998)).

5

Liposomes have been widely used to enhance the immune response. For example, a DNA vaccine constructed with the CMV promoter conjugated to env gp160 and rev genes has been shown to induce an effective immune response when inoculated via intramuscular, intraperitoneal, subcutaneous, intradermal and intranasal routes (10 *Fukushima I.N.* "Cationic liposomes are a strong adjuvant for a DNA vaccine of human immunodeficiency virus type 1" 13:1421-1428 (1997)). By immunizing with pCMV160/REV and cationic liposomes through various routes higher levels of both antibody production and delayed-type hypersensitivity were induced than by using DNA vaccine alone.

15

DNA vaccines can also be administered in combination with other agents in liposomes to increase levels of immunity. Co-administration of the DNA vaccine with IL-12 and granulocyte/macrophage CSF-expressing plasmids induced high levels of HIV-specific circulating T lymphocytes and in increase in delayed type hypersensitivity (20 when administered by the intranasal route. The results indicate that intranasal administration of this DNA vaccine with liposomes, together with IL-12 and/or granulocyte/macrophage-CSF expressing plasmids, induces a strong level of anti-HIV-1 immune response (*Okada E.* "Intranasal immunization of a DNA vaccine with IL-12 and granulocyte-macrophage colony-stimulating factor (GM-CSF)-expressing plasmids (25 in liposomes induces strong mucosal and cell-mediated immune responses against HIV-1 antigens" 159:3638-47 (1997)).

The liposome comprising a recombinant AAV virion and a biological molecule or a DNA vaccine can be delivered to a specific cell type by covalently (30 attaching a targeting moiety to a liposome or allowing the targeting moiety to become

integrated into the membrane as the liposome is formed. The targeting moiety can bind to a specific cell type, thus allowing the contents of the liposome to be delivered to a cell. For example, a targeting moiety specific for tumor cells can be incorporated into the liposome. Upon delivery of the liposome, the targeting moiety will bind to a tumor cell allowing thus allowing the toxin to enter the tumor cell. Alternatively, the targeting moiety can be a ligand that binds to a cell surface protein or receptor. Numerous cell-specific cell surface proteins are known which can be targeted by the present invention by incorporating a ligand for the cell surface protein into liposomes.

Also provided by this invention are conjugates that utilize the AAV capsid or a unique region of the AAV capsid protein (e.g. VP1, VP2 or VP3 or combinations thereof) to introduce DNA into cells. For example, the AAV VP3 protein or fragment thereof, can be conjugated to a DNA on a plasmid that is conjugated to a lipid. Cells can be infected using the targeting ability of the VP3 capsid protein to achieve the desired tissue tropism, specific to AAV. AAV VP1 and VP2 proteins can also be utilized to introduce DNA or other molecules into cells. By further incorporating the Rep protein and the AAV TRS into the DNA-containing conjugate, cells can be transduced and targeted integration can be achieved. For example, if AAV specific targeted integration is desired, a conjugate composed of the AAV VP3 capsid, AAV rep or a fragment of AAV rep, AAV TRS, the rep binding site, the heterologous DNA of interest, and a lipid, can be utilized to achieve AAV specific tropism and AAV specific targeted integration in the genome.

Further provided by this invention are chimeric viruses where AAV can be combined with herpes virus, baculovirus or other viruses to achieve a desired tropism associated with another virus. For example, the AAV ITRs could be inserted in the herpes virus and cells could be infected. Post-infection, the ITRs of AAV could be acted on by AAV rep provided in the system or in a separate vehicle to rescue AAV from the genome. Therefore, the cellular tropism of the herpes simplex virus can be combined with AAV rep mediated targeted integration. Other viruses that could be

utilized to construct chimeric viruses include, lentivirus, retrovirus, pseudotyped retroviral vectors, and adenoviral vectors.

5 In another example, AAV infects avian cells in much greater efficiencies than any other AAV. Traditionally, wild type AAV has been propagated in chicken embryonated eggs in co-infection with avian adenoviruses (i.e., Fowl adenovirus type 1, better known as CELO virus). Recently, recombinant CELO virus that can replicate in chicken embryonated eggs has been constructed (Anne-Isabelle Michou et al, 1999, J virol. 73(2): 1399). A recombinant AAV virion that encapsidates a therapeutic gene  
10 flanked by AAV ITRs can be produced in embryonated chicken eggs upon co-infection with a recombinant CELO virus expressing the AAV's rep and cap gene.

Any of the particles or virions comprising an exogenous nucleic acid encoding a protein described herein can be administered to a fertilized avian egg for the purposes  
15 of producing the recombinant protein in an avian egg. This is particularly useful for the production of vaccines as the protein produced in the avian egg can be readily purified by methods known in the art and administered to subjects in need of a vaccine.

The present invention further provides isolated nucleic acids of AAV. For  
20 example, provided is an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1 (AAV genome). This nucleic acid, or portions thereof, can be inserted into vectors, such as plasmids, yeast artificial chromosomes, or other viral vector (particle), if desired, by standard cloning methods. The present invention also provides an isolated nucleic acid consisting essentially of the nucleotide sequence set  
25 forth in SEQ ID NO:1. The nucleotides of SEQ ID NO:1 can have minor modifications and still be contemplated by the present invention. For example, modifications that do not alter the amino acid encoded by any given codon (such as by modification of the third, "wobble," position in a codon) can readily be made, and such alterations are known in the art. Furthermore, modifications that cause a resulting neutral (conserved)  
30 amino acid substitution of a similar amino acid can be made in a coding region of the

genome. Additionally, modifications as described herein for the AAV components, such as the ITRs, the p5 promoter, etc. are contemplated in this invention.

Furthermore, modifications to regions of SEQ ID NO:1 other than in the ITR, TRS Rep binding site and hairpin are likely to be tolerated without serious impact on the function of the nucleic acid as a recombinant vector.

As used herein, the term "isolated" refers to a nucleic acid separated or significantly free from at least some of the other components of the naturally occurring organism, for example, the cell structural components or viral components commonly found associated with nucleic acids in the environment of the virus and/or other nucleic acids. The isolation of the native nucleic acids can be accomplished, for example, by techniques such as cell lysis followed by phenol plus chloroform extraction, followed by ethanol precipitation of the nucleic acids. The nucleic acids of this invention can be isolated from cells according to any of many methods well known in the art.

As used herein, the term "nucleic acid" refers to single-or multiple stranded molecules which may be DNA or RNA, or any combination thereof, including modifications to those nucleic acids. The nucleic acid may represent a coding strand or its complement, or any combination thereof. Nucleic acids may be identical in sequence to the sequences which are naturally occurring for any of the novel genes discussed herein or may include alternative codons which encode the same amino acid as those provided herein, including that which is found in the naturally occurring sequence. These nucleic acids can also be modified from their typical structure. Such modifications include, but are not limited to, methylated nucleic acids, the substitution of a non-bridging oxygen on the phosphate residue with either a sulfur (yielding phosphorothioate deoxynucleotides), selenium (yielding phosphorselenoate deoxynucleotides), or methyl groups (yielding methylphosphonate deoxynucleotides).

The present invention additionally provides an isolated nucleic acid that selectively hybridizes with any nucleic acid disclosed herein, including the entire AAV genome and any unique fragment thereof, including the Rep and capsid

encoding sequences (e.g. SEQ ID NOS: 1, 2, 4, 6, 8, 10, 12, 14 and 16-24).

Specifically, the nucleic acid can selectively or specifically hybridize to an isolated nucleic acid consisting of the nucleotide sequence set forth in SEQ ID NO:1 (AAAV genome). The present invention further provides an isolated nucleic acid that

5 selectively or specifically hybridizes with an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1 (AAAV genome). By "selectively hybridizes" as used herein is meant a nucleic acid that hybridizes to one of the disclosed nucleic acids under sufficient stringency conditions without significant hybridization to a nucleic acid encoding an unrelated protein, and particularly, without  
10 detectably hybridizing to nucleic acids of AAV2 or other AAVs. Thus, a nucleic acid that selectively hybridizes with a nucleic acid of the present invention will not selectively hybridize under stringent conditions with a nucleic acid encoding a different protein or the corresponding protein from a different serotype of the virus, and vice versa. A "specifically hybridizing" nucleic acid is one that hybridizes under stringent  
15 conditions to only a nucleic acid found in AAAV. Therefore, nucleic acids for use, for example, as primers and probes to detect or amplify the target nucleic acids are contemplated herein. Nucleic acid fragments that selectively hybridize to any given nucleic acid can be used, *e.g.*, as primers and or probes for further hybridization or for amplification methods (*e.g.*, polymerase chain reaction (PCR), ligase chain reaction  
20 (LCR)). Additionally, for example, a primer or probe can be designed that selectively hybridizes with both AAAV and a gene of interest carried within the AAAV vector (*i.e.*, a chimeric nucleic acid).

Stringency of hybridization is controlled by both temperature and salt  
25 concentration of either or both of the hybridization and washing steps. Typically, the stringency of hybridization to achieve selective hybridization involves hybridization in high ionic strength solution (6X SSC or 6X SSPE) at a temperature that is about 12-25°C below the  $T_m$  (the melting temperature at which half of the molecules dissociate from their hybridization partners) followed by washing at a combination of temperature  
30 and salt concentration chosen so that the washing temperature is about 5°C to 20°C



below the  $T_m$ . The temperature and salt conditions are readily determined empirically in preliminary experiments in which samples of reference DNA immobilized on filters are hybridized to a labeled nucleic acid of interest and then washed under conditions of different stringencies. Hybridization temperatures are typically higher for DNA-RNA and RNA-RNA hybridizations. The washing temperatures can be used as described above to achieve selective stringency, as is known in the art. (Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989; Kunkel et al. *Methods Enzymol.* 1987:154:367, 1987). For the nucleic acids of the present invention, stringent hybridization conditions for a DNA:DNA hybridization can be at about 65°C (in aqueous solution) in 6X SSC or 6X SSPE followed by washing at 65°C. Therefore, the present invention provides nucleic acids that selectively hybridize to any of the nucleic acids described herein at about 65°C (in aqueous solution) in 6X SSC or 6X SSPE followed by washing at 65°C. Stringency of hybridization and washing, if desired, can be reduced accordingly as the degree of complementarity desired is decreased, and further, depending upon the G-C or A-T richness of any area wherein variability is searched for. Likewise, stringency of hybridization and washing, if desired, can be increased accordingly as homology desired is increased, and further, depending upon the G-C or A-T richness of any area wherein high homology is desired, all as known in the art.

A nucleic acid that selectively hybridizes to any portion of the AAV genome is contemplated herein. Therefore, a nucleic acid that selectively hybridizes to AAV can be of longer length than the AAV genome, it can be about the same length as the AAV genome or it can be shorter than the AAV genome. The length of the nucleic acid is limited on the shorter end of the size range only by its specificity for hybridization to AAV, *i.e.*, once it is too short, typically less than about 5 to 7 nucleotides in length, it will no longer bind specifically to AAV, but rather will hybridize to numerous background nucleic acids. Additionally contemplated by this invention is a nucleic acid that has a portion that specifically hybridizes to AAV and a portion that specifically hybridizes to a gene of interest inserted within AAV.

The present invention further provides an isolated nucleic acid encoding an avian adeno-associated virus Rep protein. The AAV Rep proteins are encoded by open reading frame (ORF) 1 of the AAV genome. Examples of the AAV Rep genes are shown in the nucleic acid set forth in SEQ ID NO:1, and include nucleic acids consisting essentially of the nucleotide sequences set forth in SEQ ID NOS:4 (Rep52), 2 (Rep78), 8 (Rep40), and 6 (Rep68), and nucleic acids comprising the nucleotide sequences set forth in SEQ ID NOS:2, 4, 6 and 8. Also contemplated herein are vectors comprising nucleotides 1-600 of SEQ ID NO: 1 which encode the first 200 amino acids of Rep. However, the present invention contemplates that the Rep nucleic acid can include any one, two, three, or four of the four Rep proteins, in any order, in such a nucleic acid. Furthermore, minor modifications are contemplated in the nucleic acid, such as silent mutations in the coding sequences, mutations that make neutral or conservative changes in the encoded amino acid sequence, and mutations in regulatory regions that do not disrupt the expression of the gene. Examples of other minor modifications are known in the art. Further modifications can be made in the nucleic acid, such as to disrupt or alter expression of one or more of the Rep proteins in order to, for example, determine the effect of such a disruption; such as to mutate one or more of the Rep proteins to determine the resulting effect, etc. However, in general, a modified nucleic acid encoding a Rep protein will have at least about 70%, about 75%, about 80%, about 85%, about 90%, about 93%, about 95%, about 98% or 100% homology to the Rep nucleic sequences described herein e.g., SEQ ID NOS: 2, 4, 6 and 8, and the Rep polypeptide encoded therein will have overall about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, about 98%, about 99% or 100% homology with the amino acid sequence described herein, e.g., SEQ ID NOS:3, 5, 7 and 9. Percent homology is determined by the techniques described herein.

The present invention also provides an isolated nucleic acid that selectively or specifically hybridizes with a nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NOS: 2, 4, 6 and 8 and an isolated nucleic acid that

selectively hybridizes with a nucleic acid comprising the nucleotide sequence set forth in SEQ ID NOS: 2, 4, 6 and 8. "Selectively hybridizing" and "stringency of hybridization" is defined elsewhere herein.

5           As described above, the present invention provides the nucleic acid encoding a Rep 40 protein and, in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO: 8, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO: 8, and a nucleic acid encoding the avian adeno-associated virus protein having the amino acid sequence set forth in SEQ  
10 ID NO: 9. The present invention also provides the nucleic acid encoding a Rep 52 protein, and in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:4, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:4, and a nucleic acid encoding the avian adeno-associated virus Rep protein having the amino acid sequence set forth in SEQ  
15 ID NO:5. The present invention further provides the nucleic acid encoding a Rep 68 protein and, in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO: 6, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO: 6, and a nucleic acid encoding the avian adeno-associated virus protein having the amino acid sequence set forth in SEQ ID NO:  
20 7. And, further, the present invention provides the nucleic acid encoding a Rep 78 protein, and in particular an isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:2, an isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:, and a nucleic acid encoding the avian adeno-associated virus Rep protein having the amino acid sequence set forth in SEQ  
25 ID NO:3. As described elsewhere herein, these nucleic acids can have minor modifications, including silent nucleotide substitutions, mutations causing conservative amino acid substitutions in the encoded proteins, and mutations in control regions that do not or minimally affect the encoded amino acid sequence.

The present invention further provides a nucleic acid encoding the entire AAV Capsid polypeptide. Furthermore, the present invention provides a nucleic acid encoding each of the three AAV coat proteins, VP1, VP2, and VP3. Thus, the present invention provides a nucleic acid encoding AAV VP1, a nucleic acid encoding  
5 AAV VP2, and a nucleic acid encoding AAV VP3. Thus, the present invention provides a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:11 (VP1); a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:13 (VP2), and a nucleic acid encoding the amino acid sequence set forth in SEQ ID NO:15 (VP3). The present invention also specifically provides a nucleic acid  
10 comprising SEQ ID NO:10 (VP1 gene); a nucleic acid comprising SEQ ID NO:12 (VP2 gene); and a nucleic acid comprising SEQ ID NO:14 (VP3 gene). The present invention also specifically provides a nucleic acid consisting essentially of SEQ ID NO:10 (VP1 gene), a nucleic acid consisting essentially of SEQ ID NO:12 (VP2 gene), and a nucleic acid consisting essentially of SEQ ID NO:14 (VP3 gene). The present  
15 invention also provides a nucleic acid comprising nucleotides 1347-2127 of SEQ ID NO: 10 (encoding amino acids 449-709 of VP1). Minor modifications in the nucleotide sequences encoding the capsid, or coat, proteins are contemplated, as described above for other AAV nucleic acids. However, in general, a modified nucleic acid encoding a capsid protein will have at least about 70%, about 75%, about 80%, about 85%, about  
20 90%, about 93%, about 95%, about 98% or 100% homology to the capsid nucleic acid sequences described herein e.g., SEQ ID NOS: 10, 12, and 14, and the capsid polypeptide encoded therein will have overall about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, about 98%, about 99% or 100% homology with the amino acid sequence described herein, e.g., SEQ ID NOS:11, 13, and 15. Nucleic acids  
25 that selectively hybridize with the nucleic acids of SEQ ID NOS: 10, 12, and 14 under the conditions described above are also provided.

The present invention also provides a cell containing one or more of the herein described nucleic acids, such as the AAV genome, AAV ORF1 and ORF2, each  
30 AAV Rep protein gene, or each AAV capsid protein gene. Such a cell can be any

desired cell and can be selected based upon the use intended. For example, cells can include bacterial cells, yeast cells, insect cells, human HeLa cells and simian Cos cells as well as other human and mammalian cells and cell lines. Primary cultures as well as established cultures and cell lines can be used. Nucleic acids of the present invention  
5 can be delivered into cells by any selected means, in particular depending upon the target cells. Many delivery means are well-known in the art. For example, electroporation, calcium phosphate precipitation, microinjection, cationic or anionic liposomes, and liposomes in combination with a nuclear localization signal peptide for delivery to the nucleus can be utilized, as is known in the art. Additionally, if the  
10 nucleic acids are in a viral particle, the cells can simply be transduced with the virion by standard means known in the art for AAV transduction. Small amounts of the recombinant AAV virus can be made to infect cells and produce more of itself.

The invention provides purified AAV polypeptides. The term "polypeptide"  
15 as used herein refers to a polymer of amino acids and includes full-length proteins and fragments thereof. Thus, "protein," polypeptide," and "peptide" are often used interchangeably herein. Substitutions can be selected by known parameters to be neutral (*see, e.g.,* Robinson WE Jr, and Mitchell WM., AIDS 4:S151-S162 (1990)). As will be appreciated by those skilled in the art, the invention also includes those  
20 polypeptides having slight variations in amino acid sequences or other properties. Such variations may arise naturally as allelic variations (*e.g.,* due to genetic polymorphism) or may be produced by human intervention (*e.g.,* by mutagenesis of cloned DNA sequences), such as induced point, deletion, insertion and substitution mutants. Minor changes in amino acid sequence are generally preferred, such as conservative amino  
25 acid replacements, small internal deletions or insertions, and additions or deletions at the ends of the molecules. Substitutions may be designed based on, for example, the model of Dayhoff, *et al.* (in *Atlas of Protein Sequence and Structure* 1978, Nat'l Biomed. Res. Found., Washington, D.C.). These modifications can result in changes in  
30 other specific mutations. The location of any modifications to the polypeptide will

often determine its impact on function. Particularly, alterations in regions non-essential to protein function will be tolerated with fewer effects on function. Elsewhere in the application regions of the AAV proteins are described to provide guidance as to where substitutions, additions or deletions can be made to minimize the likelihood of disturbing the function of the variant.

A polypeptide of the present invention can be readily obtained by any of several means. For example, the polypeptide of interest can be synthesized chemically by standard methods. Additionally, the coding regions of the genes can be recombinantly expressed and the resulting polypeptide isolated by standard methods. Furthermore, an antibody specific for the resulting polypeptide can be raised by standard methods (see, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1988), and the protein can be isolated from a cell expressing the nucleic acid encoding the polypeptide by selective hybridization with the antibody. This protein can be purified to the extent desired by standard methods of protein purification (see, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989).

Typically, to be unique, a polypeptide fragment of the present invention will be at least about 5 amino acids in length; however, unique fragments can be 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more amino acids in length. A unique polypeptide will typically comprise such a unique fragment; however, a unique polypeptide can also be determined by its overall homology. A unique polypeptide can be 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more amino acids in length. Uniqueness of a polypeptide fragment can readily be determined by standard methods such as searches of computer databases of known peptide or nucleic acid sequences or by hybridization studies to the nucleic acid encoding the protein or to the protein itself, as known in the art. The uniqueness of a polypeptide fragment can also be determined immunologically

as well as functionally. Uniqueness can be simply determined in an amino acid-by-amino acid comparison of the polypeptides.

An antigenic or immunoreactive fragment of this invention is typically an amino acid sequence of at least about 5 consecutive amino acids, and it can be derived from the AAV polypeptide amino acid sequence. An antigenic AAV fragment is any fragment unique to the AAV protein, as described herein, against which an AAV-specific antibody can be raised, by standard methods. Thus, the resulting antibody-antigen reaction should be specific for AAV.

10

The present invention provides an isolated AAV Rep protein. An AAV Rep polypeptide is encoded by ORF1 of AAV. The present invention also provides each individual AAV Rep protein. Thus the present invention provides AAV Rep 40 (e.g., SEQ ID NO: 9), or a unique fragment thereof. The present invention provides AAV Rep 52 (e.g., SEQ ID NO: 5), or a unique fragment thereof. The present invention provides AAV Rep 68 (e.g., SEQ ID NO: 7), or a unique fragment thereof. The present invention provides an example of AAV Rep 78 (e.g., SEQ ID NO: 3), or a unique fragment thereof. By "unique fragment thereof" is meant any smaller polypeptide fragment encoded by an AAV rep gene that is of sufficient length to be found only in the Rep polypeptide. Substitutions and modifications of the amino acid sequence can be made as described above and, further, can include protein processing modifications, such as glycosylation, to the polypeptide.

20

The present invention further provides an AAV Capsid polypeptide or a unique fragment thereof. AAV capsid polypeptide is encoded by ORF 2 of AAV. The present invention further provides the individual AAV capsid proteins, VP1, VP2 and VP3 or unique fragments thereof. Thus, the present invention provides an isolated polypeptide having the amino acid sequence set forth in SEQ ID NO:11 (VP1). The present invention additionally provides an isolated polypeptide having the amino acid sequence set forth in SEQ ID NO:13 (VP2). The present invention also provides

30

an isolated polypeptide having the amino acid sequence set forth in SEQ ID NO:15 (VP3). By "unique fragment thereof" is meant any smaller polypeptide fragment encoded by any AAV capsid gene that is of sufficient length to be found only in the AAV capsid protein. Substitutions and modifications of the amino acid sequence can be made as described above and, further, can include protein processing modifications, such as glycosylation, to the polypeptide. However, an AAV Capsid polypeptide including all three coat proteins will have greater than about 56% overall homology to the polypeptide encoded by the nucleotides set forth in SEQ ID NOS:10, 12 or 14. The protein can have about 65%, about 70%, about 75%, about 80%, about 85%, about 90%, 93%, 95%, 97% or even 100% homology to the amino acid sequence encoded by the nucleotides set forth in SEQ ID NOS: 10, 12 or 14. An AAV VP1 polypeptide can have at least about 58%, about 60%, about 70%, about 80%, about 90%, 93%, 95%, 97% or about 100% homology to the amino acid sequence set forth in SEQ ID NO:10. An AAV VP2 polypeptide can have at least about 58%, about 60%, about 70%, about 80%, about 90%, 93%, 95%, 97% or about 100% homology to the amino acid sequence set forth in SEQ ID NO:12. An AAV VP3 polypeptide can have at least about 60%, about 70%, about 80%, about 90%, 93%, 95%, 97% or about 100% homology to the amino acid sequence set forth in SEQ ID NO:14.

The present invention further provides an isolated antibody that specifically binds an AAV Rep protein or a unique epitope thereof. Also provided are isolated antibodies that specifically bind the AAV Rep 52 protein, the AAV Rep 40 protein, the AAV Rep 68 protein and the AAV Rep 78 protein having the amino acid sequences set forth in SEQ ID NO:5, SEQ ID NO: 9, SEQ ID NO: 7 and SEQ ID NO: 3, respectively or that specifically binds a unique fragment thereof. Clearly, any given antibody can recognize and bind one of a number of possible epitopes present in the polypeptide; thus only a unique portion of a polypeptide (having the epitope) may need to be present in an assay to determine if the antibody specifically binds the polypeptide.



The present invention additionally provides an isolated antibody that specifically binds any of the avian adeno-associated virus Capsid proteins (VP1, VP2 or VP3), a unique epitope thereof, or the polypeptide comprising all three AAV coat proteins. Also provided is an isolated antibody that specifically binds the AAV capsid protein having the amino acid sequence set forth in SEQ ID NO:11 (VP1), or that specifically binds a unique fragment thereof. The present invention further provides an isolated antibody that specifically binds the AAV Capsid protein having the amino acid sequence set forth in SEQ ID NO:13 (VP2), or that specifically binds a unique fragment thereof. The invention additionally provides an isolated antibody that specifically binds the AAV Capsid protein having the amino acid sequence set forth in SEQ ID NO:15 (VP3), or that specifically binds a unique fragment thereof. Again, any given antibody can recognize and bind one of a number of possible epitopes present in the polypeptide; thus only a unique portion of a polypeptide (having the epitope) may need to be present in an assay to determine if the antibody specifically binds the polypeptide.

The antibody can be a component of a composition that comprises an antibody that specifically binds the AAV protein. The composition can further comprise, *e.g.*, serum, serum-free medium, or a pharmaceutically acceptable carrier such as physiological saline, etc..

By "an antibody that specifically binds" an AAV polypeptide or protein is meant an antibody that selectively binds to an epitope on any portion of the AAV peptide such that the antibody binds specifically to the corresponding AAV polypeptide without significant background. Specific binding by an antibody further means that the antibody can be used to selectively remove the target polypeptide from a sample comprising the polypeptide or and can readily be determined by radioimmunoassay (RIA), bioassay, or enzyme-linked immunosorbant (ELISA) technology. An ELISA method effective for the detection of the specific antibody-antigen binding can, for example, be as follows: (1) bind the antibody to a substrate;

(2) contact the bound antibody with a sample containing the antigen; (3) contact the above with a secondary antibody bound to a detectable moiety (e.g., horseradish peroxidase enzyme or alkaline phosphatase enzyme); (4) contact the above with the substrate for the enzyme; (5) contact the above with a color reagent; (6) observe the color change.

An antibody can include antibody fragments such as Fab fragments which retain the binding activity. Antibodies can be made as described in, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1988). Briefly, purified antigen can be injected into an animal in an amount and in intervals sufficient to elicit an immune response. Antibodies can either be purified directly, or spleen cells can be obtained from the animal. The cells are then fused with an immortal cell line and screened for antibody secretion. Individual hybridomas are then propagated as individual clones serving as a source for a particular monoclonal antibody.

The present invention additionally provides a method of screening a cell for infectivity by AAV comprising contacting the cell with AAV and detecting the presence of AAV in the cells. AAV particles can be detected using any standard physical or biochemical methods.

The present invention provides a method of screening for cells that are permissive to AAV infection comprising identifying the presence of N-linked terminal lactose on the surface of a cell, contacting the N-linked terminal lactose containing cell with AAV and detecting the presence of AAV virus in the cell, whereby if AAV virus is detected in the cells, the N-linked terminal lactose containing cell is permissive to AAV infection. In one example of such a method, based on the teaching in the Examples, uses *Erythrina corralodendron* lectin to detect cells that would allow efficient binding of AAV and possible transduction. A closely related method using sialic acid binding lectins to screen for AAV5 transduction is

described in Walters et al. (Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity, J Virol. 2001 Aug; 75(15):6884-93, incorporated herein by reference).

5

For the screening methods of the present invention, monoclonal antibodies to different forms of conjugated lactose can be produced (Sato et al JBC 2000. May 19;275(20):15422-31). Cells are contacted with these antibodies to select cells that contain the appropriate N-linked terminal lactose. A number of antibodies exist which bind specific lactose conjugates and can be used to screen for N-linked terminal lactose containing cells. These antibodies can be fluorescently labeled and used in situ. Alternatively, antibodies can be bound to a plate and target cells added. The wells are then washed and cells that express the antigen will bind to the N-linked terminal lactose antibody. Cells that bind to the sialic acid can be visualized by staining. Another way to screen for permissive cells is to chemically remove the glycans from the cell surface and fractionate these by thin layer chromatography. The presence of the correct form of N-linked terminal lactose can be confirmed by hybridizing the blot with labeled virus. Free virus is washed off and the specifically bound virus visualized by detecting the label. Alternatively, whole membrane proteins could be used and separated by PAGE, transferred to a membrane and probed as described above.

Additionally, physical methods that can be used for this detection include DNA based methods such as 1) polymerase chain reaction (PCR) for viral DNA or RNA or 2) direct hybridization with labeled probes, and immunological methods such as by 3) antibody directed against the viral structural or non- structural proteins. Catalytic methods of viral detection include, but are not limited to, detection of site and strand specific DNA nicking activity of Rep proteins or replication of an AAV origin-containing substrate. Reporter genes can also be utilized to detect cells that transduce AAV. For example,  $\beta$ -gal, green fluorescent protein or luciferase can be inserted into a recombinant AAV. The cell can then be contacted with the recombinant AAV,

30

either *in vitro* or *in vivo* and a colorimetric assay could detect a color change in the cells that would indicate transduction of AAV in the cell. Additional detection methods are outlined in Fields, *Virology*, Raven Press, New York, New York. 1996.

5           For screening a cell for infectivity by AAV, wherein the presence of AAV in the cells is determined by nucleic acid hybridization methods, a nucleic acid probe for such detection can comprise, for example, a unique fragment of any of the AAV nucleic acids provided herein. The uniqueness of any nucleic acid probe can readily be determined as described herein. Additionally, the presence of AAV in cells can be  
10 determined by fluorescence, antibodies to gene products, focus forming assays, plaque lifts, Western blots and chromogenic assays. The nucleic acid can be, for example, the nucleic acid whose nucleotide sequence is set forth in SEQ ID NO: 1, 2, 4, 6, 8, 10, 12, 14, and 16-24 or a unique fragment thereof.

15           The present invention includes a method of determining the suitability of an AAV vector for administration to a subject comprising administering to an antibody-containing sample from the subject an antigenic fragment of an isolated AAV Rep or Capsid protein, and detecting neutralizing antibody-antigen reaction in the sample, the presence of a neutralizing reaction indicating the AAV vector may be unsuitable for  
20 use in the subject. The present method of determining the suitability of an AAV vector for administration to a subject can comprise contacting an antibody-containing sample from the subject with a unique antigenic or immunogenic fragment of an AAV Rep protein (e.g. Rep 40, Rep 52, Rep 68, Rep 78) and detecting an antibody-antigen reaction in the sample, the presence of a reaction indicating the AAV vector  
25 to be unsuitable for use in the subject. The AAV Rep proteins are provided herein, and their antigenic fragments are routinely determined. The AAV capsid protein can be used to select an antigenic or immunogenic fragment, for example from the amino acid sequence set forth in SEQ ID NO:11 (VP1), the amino acid sequence set forth in SEQ ID NO: 13 (VP2) or the amino acid sequence set forth in SEQ ID NO:15 (VP3).  
30 Alternatively, or additionally, an antigenic or immunogenic fragment of an isolated

AAAV Rep protein can be utilized in this determination method. The AAAV Rep protein from which an antigenic fragment is selected can have the amino acid sequence encoded by the nucleic acid set forth in SEQ ID NO:1, the amino acid sequence set forth in SEQ ID NO:3, or the amino acid sequence set forth in SEQ ID NO:5, the  
5 amino acid sequence set forth in SEQ ID NO: 7, or the amino acid sequence set forth in SEQ ID NO:9.

The AAAV polypeptide fragments can be analyzed to determine their antigenicity, immunogenicity and/or specificity. Briefly, various concentrations of a  
10 putative immunogenically specific fragment are prepared and administered to a subject and the immunological response (e.g., the production of antibodies or cell mediated immunity) of an animal to each concentration is determined. The amounts of antigen administered depend on the subject, e.g. a human, rabbit or a guinea pig, the condition of the subject, the size of the subject, etc. Thereafter an animal so inoculated with the  
15 antigen can be exposed to the AAAV viral particle or AAAV protein to test the immunoreactivity or the antigenicity of the specific immunogenic fragment. The specificity of a putative antigenic or immunogenic fragment can be ascertained by testing sera, other fluids or lymphocytes from the inoculated animal for cross reactivity with other closely related viruses, such as AAV1-8.

20

By the "suitability of an AAAV vector for administration to a subject" is meant a determination of whether the AAAV vector will elicit a neutralizing immune response upon administration to a particular subject. A vector that does not elicit a significant immune response is a potentially suitable vector, whereas a vector that elicits a  
25 significant, neutralizing immune response (e.g. at least 90%) is thus likely to be unsuitable for use in that subject. Significance of any detectable immune response is a standard parameter understood by the skilled artisan in the field. For example, one can incubate the subject's serum with the virus, then determine whether that virus retains its ability to transduce cells in culture. If such virus cannot transduce cells in culture, the  
30 vector likely has elicited a significant immune response.

Alternatively, or additionally, one skilled in the art could determine whether or not AAV administration would be suitable for a particular cell type of a subject. For example, the artisan could culture muscle cells *in vitro* and transduce the cells with AAV in the presence or absence of the subject's serum. If there is a reduction in transduction efficiency, this could indicate the presence of a neutralizing antibody or other factors that may inhibit transduction. Normally, greater than 90% inhibition would have to be observed in order to rule out the use of AAV as a vector. However, this limitation could be overcome by treating the subject with an immunosuppressant that could block the factors inhibiting transduction.

As will be recognized by those skilled in the art, numerous types of immunoassays are available for use in the present invention to detect binding between an antibody and an AAV polypeptide of this invention. For instance, direct and indirect binding assays, competitive assays, sandwich assays, and the like, as are generally described in, *e.g.*, U.S. Pat. Nos. 4,642,285; 4,376,110; 4,016,043; 3,879,262; 3,852,157; 3,850,752; 3,839,153; 3,791,932; and Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, N.Y. (1988). For example, enzyme immunoassays such as immunofluorescence assays (IFA), enzyme linked immunosorbent assays (ELISA) and immunoblotting can be readily adapted to accomplish the detection of the antibody. An ELISA method effective for the detection of the antibody bound to the antigen can, for example, be as follows: (1) bind the antigen to a substrate; (2) contact the bound antigen with a fluid or tissue sample containing the antibody; (3) contact the above with a secondary antibody specific for the antigen and bound to a detectable moiety (*e.g.*, horseradish peroxidase enzyme or alkaline phosphatase enzyme); (4) contact the above with the substrate for the enzyme; (5) contact the above with a color reagent; (6) observe color change.

The antibody-containing sample of this method can comprise any biological sample which would contain the antibody or a cell containing the antibody, such as blood, plasma, serum, bone marrow, saliva, urine and mucus.

5           The present invention also provides a method of producing the AAV virus by transducing a cell with the nucleic acid encoding the virus. The present invention also provides AAV produced by the method of transducing a cell with the nucleic acid encoding the virus.

10           The present method further provides a method of delivering an exogenous (heterologous) nucleic acid to a cell comprising administering to the cell an AAV particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to the cell.

15           The AAV ITRs in the vector for the herein described delivery methods can be AAV ITRs (SEQ ID NOS: 16 and 17). Furthermore, the AAV ITRs in the vector for the herein described nucleic acid delivery methods can also comprise AAV1-8 inverted terminal repeats.

20           The present invention also includes a method of delivering a heterologous nucleic acid to a subject comprising administering to a cell from the subject an AAV particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject. The AAV ITRs can be any AAV ITRs, including AAV  
25 ITRs, AAV5 ITRs and AAV2 ITRs. For example, in an *ex vivo* administration, cells are isolated from a subject by standard means according to the cell type and placed in appropriate culture medium, again according to cell type (*see, e.g.*, ATCC catalog). Viral particles are then contacted with the cells as described above, and the virus is allowed to transduce the cells. Cells can then be transplanted back into the subject's  
30 body, again by means standard for the cell type and tissue (*e. g.*, in general, U.S. Patent

No. 5,399,346; for neural cells, Dunnett, S.B. and Björklund, A., eds., *Transplantation: Neural Transplantation-A Practical Approach*, Oxford University Press, Oxford (1992)). If desired, prior to transplantation, the cells can be studied for degree of transduction by the virus, by known detection means and as described herein. Cells for  
5 *ex vivo* transduction followed by transplantation into a subject can be selected from those listed above, or can be any other selected cell. Preferably, a selected cell type is examined for its capability to be transfected by AAV. Preferably, the selected cell will be a cell readily transduced with AAV particles; however, depending upon the application, even cells with relatively low transduction efficiencies can be useful,  
10 particularly if the cell is from a tissue or organ in which even production of a small amount of the protein or antisense RNA encoded by the vector will be beneficial to the subject.

The present invention further provides a method of delivering a nucleic acid to a  
15 cell in a subject comprising administering to the subject an AAV particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to a cell in the subject. Administration can be an *ex vivo* administration directly to a cell removed from a subject, such as any of the cells listed above, followed by replacement of the cell back into the subject, or  
20 administration can be *in vivo* administration to a cell in the subject. For *ex vivo* administration, cells are isolated from a subject by standard means according to the cell type and placed in appropriate culture medium, again according to cell type (*see, e.g.*, ATCC catalog). Viral particles are then contacted with the cells as described above, and the virus is allowed to transfect the cells. Cells can then be transplanted back into  
25 the subject's body, again by means standard for the cell type and tissue (*e. g.*, for neural cells, Dunnett, S.B. and Björklund, A., eds., *Transplantation: Neural Transplantation-A Practical Approach*, Oxford University Press, Oxford (1992)). If desired, prior to transplantation, the cells can be studied for degree of transfection by the virus, by known detection means and as described herein.

30



The present invention further provides a method of delivering a nucleic acid to a cell in a subject having neutralizing antibodies to AAV1-8 comprising administering to the subject an AAV particle containing a vector comprising the nucleic acid, thereby delivering the nucleic acid to a cell in the subject. A subject that has neutralizing  
5 antibodies to AAV1-8 can readily be determined by any of several known means, such as contacting AAV1-8 protein(s) with an antibody-containing sample, such as blood, from a subject and detecting an antigen-antibody reaction in the sample. Delivery of the AAV1-8 particle can be by either *ex vivo* or *in vivo* administration as herein described. Thus, a subject who might have an adverse immunogenic reaction to a  
10 vector administered in an AAV2 viral particle can have a desired nucleic acid delivered using an AAV1-8 particle. This delivery system can be particularly useful for subjects who have received therapy utilizing AAV1-8 particles in the past and have developed antibodies to AAV1-8. An AAV regimen can now be substituted to deliver the desired nucleic acid.

15

In any of the methods of delivering heterologous nucleic acids to a cell or subject described herein, the AAV-conjugated nucleic acid or AAV particle-conjugated nucleic acids described herein can be used.

20 *In vivo* administration to a human subject or an animal model can be by any of many standard means for administering viruses, depending upon the target organ, tissue or cell. Virus particles can be administered orally, parenterally (*e.g.*, intravenously), by intramuscular injection, by direct tissue or organ injection, by intraperitoneal injection, topically, transdermally, via aerosol delivery, via the mucosa or the like. Viral nucleic  
25 acids (non-encapsidated) can also be administered, *e.g.*, as a complex with cationic liposomes, or encapsulated in anionic liposomes. The present compositions can include various amounts of the selected viral particle or non-encapsidated viral nucleic acid in combination with a pharmaceutically acceptable carrier and, in addition, if desired, may include other medicinal agents, pharmaceutical agents, carriers, adjuvants, diluents, etc.  
30 Parental administration, if used, is generally characterized by injection. Injectables can

be prepared in conventional forms, either as liquid solutions or suspensions, solid forms suitable for solution or suspension in liquid prior to injection, or as emulsions. Dosages will depend upon the mode of administration, the disease or condition to be treated, and the individual subject's condition, but will be that dosage typical for and used in  
5 administration of other AAV vectors, such as AAV2 vectors. Often a single dose can be sufficient; however, the dose can be repeated if desirable.

Administration methods can be used to treat brain disorders such as Parkinson's disease, Alzheimer's disease, and demyelination disease. Other diseases that can be  
10 treated by these methods include metabolic disorders such as , musculoskeletal diseases, cardiovascular disease, cancer, and autoimmune disorders.

Administration of this recombinant AAV virion to the cell can be accomplished by any means, including simply contacting the particle, optionally  
15 contained in a desired liquid such as tissue culture medium, or a buffered saline solution, with the cells. The virion can be allowed to remain in contact with the cells for any desired length of time, and typically the virion is administered and allowed to remain indefinitely. For such *in vitro* methods, the virion can be administered to the cell by standard viral transduction methods, as known in the art and as exemplified  
20 herein. Titers of virus to administer can vary, particularly depending upon the cell type, but will be typical of that used for AAV transduction in general which is well known in the art. Additionally the titers used to transduce the particular cells in the present examples can be utilized.

25 The cells that can be transduced by the present recombinant AAV virion can include any desired cell, such as the following cells and cells derived from the following tissues, human as well as other mammalian tissues, such as primate, horse, sheep, goat, pig, dog, rat, and mouse and avian species: Adipocytes, Adenocyte, Adrenal cortex, Amnion, Aorta, Ascites, Astrocyte, Bladder, Bone, Bone marrow,  
30 Brain, Breast, Bronchus, Cardiac muscle, Cecum, Cervix, Chorion, Colon, Conjunctiva,

Connective tissue, Cornea, Dermis, Duodenum, Endometrium, Endothelium,  
 Endothelial cells, Epithelial tissue, Epithelial cells, Epidermis, Esophagus, Eye, Fascia,  
 Fibroblasts, Foreskin, Gastric, Glial cells, Glioblast, Gonad, Hepatic cells, Histocyte,  
 Ileum, Intestine, small Intestine, Jejunum, Keratinocytes, Kidney, Larynx, Leukocytes,  
 5 Lipocyte, Liver, Lung, Lymph node, Lymphoblast, Lymphocytes, Macrophages,  
 Mammary alveolar nodule, Mammary gland, Mastocyte, Maxilla, Melanocytes,  
 Mesenchymal, Monocytes, Mouth, Myelin, Myoblasts Nervous tissue, Neuroblast,  
 Neurons, Neuroglia, Osteoblasts, Osteogenic cells, Ovary, Palate, Pancreas, Papilloma,  
 Peritoneum, Pituicytes, Pharynx, Placenta, Plasma cells, Pleura, Prostate, Rectum,  
 10 Salivary gland, Skeletal muscle, Skin, Smooth muscle, Somatic, Spleen, Squamous,  
 Stomach, Submandibular gland, Submaxillary gland, Synoviocytes, Testis, Thymus,  
 Thyroid, Trabeculae, Trachea, Turbinate, Umbilical cord, Ureter, and Uterus.

The methods of the present invention are also useful for the delivery of AAV  
 15 vectors that express ribozymes or small interfering RNAs (siRNAs). Both methods can  
 reduce protein expression by minimizing or completely abolishing mRNA levels of  
 targeted genes. Applications in the poultry industry are also contemplated. These  
 include delivery of a ribozyme or siRNA against chicken myostatin, a gene controlling  
 muscle mass. In addition, the ability of AAV vectors to deliver genes to a variety of  
 20 tissues to express genetic information effectively for long periods of time, and to have a  
 good safety profile make avian AAVs an attractive vector for genetic immunization of  
 chickens. Avian AAV vectors could be used for *in ovo* or post-hatch vaccination of  
 chickens against diseases such as Marek's, coccidiosis, Newcastle disease, etc.

25 Also, a hallmark of avian AAV infection is the inhibition of avian viruses  
 during co-infection. The present invention shows that this is a function of the avian  
 AAV non-structural proteins. Incorporation of these sequences into a viral vector or  
 addition of recombinant protein to eggs could be used as a method to inhibit viral  
 infection and promote growth/development.

30

A method of blocking AAV infection is provided. The method is based on the findings in the Examples that AAV requires N-linked terminal lactose present on cell surface proteins for efficient binding and entry. Thus, lactose conjugates, dendrimer nanoparticles with terminal lactose, or *Erythrina corralodendron* lectin can be used as  
5 agents to block AAV infection of a cell. The synthesis of dendrimers has been described (Schchepinov, M.S., Udalova, I.A., Bridgman, A.J., Southern, E.M., 1997, Nucleic Acids Res. 25:4447-4454).

A method of inducing an immune response to AAV in a subject comprising administering an AAV particle comprising the capsid protein (SEQ ID NO: 11, SEQ ID NO: 12 or SEQ ID NO:13) or epitope thereof, wherein the capsid protein or epitope thereof comprises an epitope that induces an immune response in a subject. The capsid protein can also include epitopes of other (non-AAV) proteins (as described herein) such that an immune response is directed against the non-AAV epitope.

A method of blocking an immune response against AAV in a subject comprising administering an AAV particle comprising the capsid protein (SEQ ID NO: 11, SEQ ID NO: 12 or SEQ ID NO:13) or epitope thereof, wherein the capsid protein or epitope thereof comprises an epitope that blocks the immune response to AAV in a subject. The capsid protein can also include epitopes of other (non-AAV) proteins (as described herein) such that the immune response that is blocked is an immune response directed against the non-AAV epitope.

A method of producing a recombinant protein is provided, comprising administering an AAV particle comprising an exogenous nucleic acid encoding a protein to an embryonated avian egg; and b) purifying the protein from the egg. The protein purified by a method of the invention is also provided.

## EXAMPLES

To understand the nature of AAV virus and to determine its usefulness as a vector for gene transfer, it was cloned and sequenced.

5

### Materials and Methods

**Cell culture and virus propagation.** 293T and COS cells were maintained in IMEM and AMEM, respectively, containing 10% FBS, DF1 cells (spontaneously immortalized chicken embryonic fibroblasts), QNR cells (quail neuroretinal cells), A549 and primary chicken embryonic fibroblasts (CEF) were maintained in DMEM supplemented with 10% FBS, primary chicken embryonic kidney cells were maintained in BME supplemented with 10% FBS, primary chicken pituitary cells were maintained in DMEM supplemented with 5% horse serum, QT6 cells (quail fibrosarcoma) were maintained in Ham's F12K supplemented with 10% FBS, LMH cells (chicken hepatoma cells) cells were maintained in Waymouth's media supplemented with 10% FBS, DT-90 (chicken lymphoblastoma cells) were maintained in DMEM supplemented with 15% FBS, 5% chicken serum and 0.015%  $\beta$ -mercaptoethanol. Human primary fibroblasts were obtained from Clonetics and maintained in serum-free proprietary medium supplied by the manufacturer. AAV (ATCC, VR-865) was propagated in ten day old Spafas pathogen free embryonated chick eggs co-infected with the Phelps strain of fowl adenovirus type I (FAV1; ATCC, VR-486). AAV at  $10^4$ - $10^7$  and FAV1 at  $10^5$  infectious particles in saline were simultaneously injected in the chorioallantoic cavity of eggs and incubated for 96 hrs at 37 C. At the end of the incubation allantoamniotic fluids (AAFs) were harvested and clarified by centrifugation at 6000g for 10 min.

**Viral DNA isolation, cloning and sequencing.** Virus from infected clarified AAFs was precipitated by centrifugation at 100,000 g for 2 hr. The supernatant was discharged and the virus-containing pellet was resuspended in proteinase K digestion buffer (50 mM Tris pH=8, 20mM EDTA, 0.5% SDS, 200  $\mu$ g/ml proteinase K) and

incubated at 45 C for 2 h. Following a phenol-chloroform extraction and ethanol precipitation, the viral DNA was resuspended in TE buffer containing 0.1 M NaCl. The single stranded viral DNA was annealed by heating to 95 C for 5 min followed by slow cooling to 65 C for 6 h. The annealed viral DNA was separated electrophoretically in 1% agarose gel and the double stranded AAV DNA of approximately 4.7 kb was excised and purified using a gel extraction kit (Qiagen). The viral DNA was further processed to fill in the ends by treating with DNA polymerase (Klenow fragment) at 37 C for 15 min in the presence of dNTPs. The whole genome was then blunt end cloned in the pPCR-script cloning vector containing the LacZ gene allowing blue-white screening of ampicillin resistant colonies (Stratagene). Colonies that contained large inserts (4.7kb) were initially screened by restriction digestion and three clones were selected for sequencing. No sequence differences were found in these three clones. Sequence of the entire genome (except ITRs) was determined using an ABI 373A automated sequencer and FS dye-terminator chemistry (ABI). Due to high degree of secondary structure, ITRs were sequenced by isothermal non-cycling sequencing chemistry using radiolabeled dCTP (Epicentre). One of the clones (pAAAV) that contained the entire consensus sequence of AAV was further used to generate packaging and vector plasmids for construction of recombinant AAV (rAAAV) virus. The complete DNA sequence of AAV have been submitted to GenBank (Accession number AY186198 ).

**Sequence analysis.** DNA and protein sequence alignments were performed using the Clustal W multiple sequence alignment tool of the Biology Workbench web based software (SDSC). Promoters, transcription initiation and splice sites were predicted using the Neural Network Promoter Prediction web paged software (BDGP). The presence of potential transcription binding sites was analyzed using the MatInspector computer program (54). Putative motifs in the Rep proteins were identified using the BLIMPS program that search for motifs in the Blocks protein database (28).

**Southern blot hybridization.** The ability of pAAAV to support self-excision,

packaging and generation of nuclease resistant wild type AAV particles was examined. 293T cells seeded in 6-well plates were transfected using calcium phosphate co-precipitation with pAAV alone, pAAV plus pAd12 (a helper plasmid containing the E2 and E4 ORFs and VA RNAs of Ad5) and pAAV plus infection with Ad5. In addition, LMH cells seeded in gelatin-coated 6-well plates were similarly transfected with pAAV alone or with pAAV plus infection with FAV1. After 48 hr, clarified lysates were prepared using three freeze-thaw cycles and centrifugation at 3800xg for 20 min. The lysate (~100 µl) was treated with 5 units of DNase for 2 hr at 37 to remove vector and unpackaged progeny. Subsequently, the solution were adjusted to contain 20 mM EDTA (pH=8), 0.5 % SDS and 200 µg/ml proteinase K and incubated at 45 C for 2 hr. After one phenol-chloroform extraction, nucleic acids were precipitated with addition of an equal volume of isopropanol, and the pellet was resuspended in 30 µl of TE buffer containing 0.1 M NaCl. The samples were heated to 95 C for 5 min, slowly cooled down to 65 C at which point and incubated for 5hr. After electrophoresis and blotting, the membrane was probed with a 32P-labeled 1.2 kb BamH1 fragment of pAAV.

**Generation of recombinant AAV particles.** For production of recombinant particles we three different helper plasmids were generated and examined, pMA<sub>3</sub>VRC, pCA<sub>3</sub>VRC, pA<sub>3</sub>VRC, containing the AAV *rep* and *cap* genes under control of an MMTV, CMV or the native p5 promoters, respectively. For generation of pMA<sub>3</sub>VRC, the *rep* and *cap* ORFs (nucleotides 243-4482) was produced by PCR with pfu polymerase (Stratagene) as specified by the manufacturer using primers containing BstZ107 and NotI sites. The PCR products were digested with BstZ107 and NotI and ligated in a BstZ107/NotI fragment of pMMTV2.1 (18) containing an MMTV promoter and SV40 polyA. For generation of pCA<sub>3</sub>VRC, the *rep* and *cap* ORFs (nucleotides 243-4482) was produced by PCR with pfu polymerase and blunt end ligated in the pCMV- script (Stratagene) vector, which contains a CMV promoter and SV40 polyA. For generation of pA<sub>3</sub>VRC, the *rep* and *cap* genes of AAV including the p5 promoter and polyA signal (nucleotides 142-4516) was produced by PCR using pfu polymerase

and blunt-end ligated in pPCR-script. Orientation of inserts was verified by restriction digestion analysis, and final clones confirmed by sequencing. For generation of the vector carrying the  $\beta$ -galactosidase gene flanked by AAV ITRs, the plasmid pAAV was digested with BsmB1 (NEB). BsmB1 does not cut in the plasmid backbone but cut at positions 838, 1111, 2590, 4419 and 4530 of the AAV genome. The resulting fragment that contained the plasmid backbone and 700 bp of AAV genome flanked by ITRs was used to ligate a BsmB1-BsmI linker. The resulting plasmid was digested with Pml1 (cuts at nucleotide 146 of AAV genome) and BsmI and used to ligate a BstZ107-BsmI fragment of pAAV<sub>2</sub>RnLacZ (18) that contains the  $\beta$ -galactosidase gene under control of an RSV promoter and SV40 polyA tail. The resulting plasmid (pA<sub>3</sub>VRSV $\beta$ Gal) was co-transfected with one of the helper plasmids described above and pAd12 in 293T cells plated in 150 cm dishes. Forty-eight hours post-transfection, cells were harvested and quantitated with a hemacytometer, and rAAV prepared using standard CsCl gradient purification. The number of rAAV genomes was estimated using real time quantitative PCR (QPCR) and expressed as nuclease resistant particles per cell recovered after transfections (DRP/cell). Titration of rAAV was performed in exponentially growing CEF, DF-1, LMH, QNR, QT6, DT-90, 293T, COS and primary embryonic chicken kidney cells and non-dividing primary pituitary cells plated in 96 well plates, and transduced with serial dilutions of recombinant virus for 48 h as previously described (20).

To obtain AAV genomic DNA for cloning, a stock of AAV was obtained from ATCC (VR-865) and coinfecting with Fowl adenovirus type I in day 10 embryonated chicken eggs. Virus was concentrated after subjecting infected allantoamniotic fluids to high-speed centrifugation. Viral DNA was released by SDS-Proteinase K digestion and purified by gel electrophoresis after annealing the complementary single strands by heating the purified DNA to 95°C and slowly cooling to 65°C. Preliminary experiments indicated that 10<sup>5</sup> infectious particles of FAV1 resulted in productive infection without succumbing the embryo prematurely. Coinfection with at least 10<sup>5</sup> infectious particles of AAV was required to detect viral



DNA (~4.7 kb) by ethidium bromide staining. After recovery and end-filling, the double stranded AAV genome was blunt-end ligated and cloned into pPCR-script. Several clones that contained an insert of approximately 4.7 kb were initially screened by restriction digestion and all gave bands similar in size to those previously reported  
5 (30). Three of these clones were sequenced and all gave identical sequences. One of the clones was randomly selected and used in subsequent analysis (pAAAV).

To verify that pAAAV can support self-excision, viral DNA replication, and packaging in mammalian and avian cells, viral lysates were prepared from 293T and  
10 LMH cells transfected with pAAAV and infected with wild type Ad5 or FAV1, respectively. In addition, the ability of an Ad5 plasmid to provide helper functions was examined in 293T cells. Southern blot analysis showed encapsidated (nuclease resistant particles) AAV progeny in the presence of wtAd5 or Ad helper plasmid in 293T cells and FAV1 in LMH cells but not in the absence (Fig. 1a and b). This result  
15 suggests that pAAAV can support rescue of AAV in mammalian and avian cells in the presence of mammalian or avian adenoviral genes.

The AAV ITR is composed of 142 nucleotides with the first 122 forming the characteristic T-shaped palindromic structure (Fig 3), and it is 60-62% homologous  
20 with the ITRs of serotypes 2, 3, 4, and 6 and 48% homologous with AAV5. A tandem repeat of GAGY in the ITR, which serves as the binding element of Rep78 and Rep68 (RBE), is conserved between AAV and the other AAVs (Fig. 3,4). The *trs* recognition motif of serotypes 2, 3, 4 and 6 (CCGGT' TG) is highly homologous with that of the putative AAV *trs* (CCGGT'CG) and weekly homologous with AAV5 *trs*  
25 site (ACGGT'GT). In addition, the spacing between the RBE and the putative *trs* is similar to that found in other serotypes, a characteristic that has been shown to be essential for Rep activity (12).

It has been proposed that a potential inverted repeat flanking the core *trs*  
30 sequence of AAV serotypes might be required for Rep *trs* nicking (11). Such an

inverted repeat is not found around the AAV *trs* sequence. This observation may indicate that Avian Rep nicking does not require any secondary structure around the core *trs* element. Methylation interference experiments have indicated the importance of the CTTTG motif found at the tip of one palindrome in AAV2 Rep binding (57).

- 5 Most of this motif is conserved in AAV ITR (CTTCG) and only one T residues is changed to C. Interestingly, the AAV4 ITR has a similar substitution in this motif (CTCTG). Thus, irrespectively of the overall nucleotide sequence homology, the secondary structure and the elements required for viral replication are conserved in the AAV ITR.

10

- The entire AAV genome (Fig. 3) is 4,694 nucleotides in length and has similar organization with that of other AAVs. It has two inverted terminal repeats and two distinct ORFs. The entire genome of AAV displays 56-65% identity at the nucleotide level with the other known AAVs. The p5 promoter region of AAV is much shorter and shows some divergence from homologous regions of other AAV serotypes. Core regulatory elements such as the TATAA box and Ebox/USF are conserved, however YY1 and Rep binding sites are not present. This suggests that AAV gene expression might be regulated differently from that of other AAVs. The p19 promoter, the p40 promoter, and poly(A) can also be identified in the AAV genome by homology to those in primate AAV serotypes. Based on the general organization and sequence, these elements are highly conserved.

- Clustal W protein sequence alignment indicate the left ORF of AAV is 46-54% identical and equally divergent from that of the primate AAVs and the GPV Rep ORF (Fig. 4 a) and only 18-22% identical with the Rep ORF of other mammalian autonomous parvovirus. In comparison, the Rep ORF of isolates 1, 2, 3, 4, 6, 7 and 8 are greater than 90% similar and approximately 67-70% identical with that of AAV5 Rep ORF. The central region of the AAV Rep ORF (aa 322 to 470), which is present in all Rep proteins, displays the greatest identity (82%) with the same region of the other AAVs and the GPV. This region of the Rep proteins is necessary for ATPase and

30

helicase activity and contains an ATP-binding site (aa 334 to 349) and a divalent cation binding site at amino acid residue 421 (44, 61, 65). The amino terminus (aa 1 to 251) is 42-45% similar between AAV and the other AAVs. This region of the Rep78 and Rep68 proteins is required for DNA binding and *trs* endonuclease activities (22, 50). A  
5 tyrosine residue at 155 is homologous to the Tyr156 in AAV2 that functions as the active nucleophile in the *trs* endonuclease site (22, 62). The active site is assembled by the spatial convergence of a divalent metal ion that is tetrahedrally coordinated by Asp24, Glu83, His90 and His92. In addition Glu6 is required for the correct orientation of the two active sites imidazoles from His90 and His92 (31). All of these amino acid  
10 residues are strictly conserved among AAV serotypes including AAV. Furthermore, a helix region important for Rep multimerization (aa 159-179) is also conserved in AAV. The carboxyl terminal portion (aa 490-662) of the unspliced AAV Rep proteins appears highly divergent, displaying less than 15% homology with the primate serotypes. However, a characteristic Zinc finger motif was identified using the  
15 BLIMPS algorithm. This feature is conserved in all AAV serotypes.

The right ORF of AAV, which encodes the three viral capsid proteins, is approximately 54-57% identical to the capsid ORF of the other AAVs and the GPV (Fig. 4 b). It has been previously reported (6) that the AAV capsid proteins VP1,  
20 VP2 and VP3 have apparent molecular weights of 92, 69 and 61 kDa, respectively, as determined by SDS-PAGE. The calculated molecular masses based on amino acid composition for VP1, VP2 and VP3 are 83, 67 and 60 kDa. We also subjected purified AAV virions to SDS-PAGE and found that they have MW 91, 68 and 60 kDa (data not shown). As with the primate AAVs and the goose and duck autonomous  
25 parvovirus, the AAV cap gene contains two ATG initiator codons, one for VP1 and the other for VP3. The unusual ACG initiator codon for VP2 is also conserved in AAV.

Clustal W alignment of the VP ORFs indicated the presence of conserved and  
30 divergent regions. The N terminus of VP1 (aa 1-143), which is required for particle

formation, is relatively conserved among AAV, AAV2, AAV4, AAV5 and GPV. However, the start site for VP2 and VP3 are found in a divergent region. Based on the published three-dimensional structure of the canine parvovirus and comparisons of parvovirus capsid sequences (15), most of the divergent regions among AAV, AAV2, AAV4 and AAV5 and GPV are located on the exterior of the virus, thus suggesting different uptake mechanisms and altered tissue tropism.

In the present study, recombinant AAV particles containing the gene for nuclear localized  $\beta$ -galactosidase were constructed. Virus was produced as previously described (19, 20) by constructing a vector plasmid containing the  $\beta$ -galactosidase gene under control of an RSV promoter flanked by AAV ITRs (pA3V $\beta$ gal, Fig. 5 a), and a helper plasmid containing the AAV rep and cap genes. Virus was isolated from 293T cell lysates by CsCl banding, and the distribution of recombinant virus across the gradient was determined by QPCR analysis of gradient fractions. The majority of packaged genomes were found in fractions with a density of 1.42 g/cm<sup>3</sup>, which is similar to that of wt AAV. We also examined the yield of rAAV when using helper plasmids with the rep gene under control of three different promoters, CMV, MMTV or the native P5 promoter (Fig. 5 a). The different helper plasmids (pCA3VRC, pMA3VRC, pA3VRC) were co-transfected with pA3V $\beta$ gal and an adenovirus helper plasmid in 293T cells and rAAV was purified from the three different CVLs using CsCl gradients. The number of rAAV genomes was determined by QPCR. In three independent trials, the yield of rAAV was 5-fold and 15-fold greater using the stronger CMV promoter compared with the MMTV and the native P5 promoter, respectively (Fig 5 a). This finding with rAAV is in contrast to previous work with AAV2 that demonstrated the use of a CMV promoter inhibited the production of rAAV2 (39).

In preliminary studies, it was observed that the addition of detergents during virus purification affected infectivity. To better understand the effect of detergents, we prepared rAAV in the presence of the following conditions: 0.5% deoxycholate, 0.5%

CHAPS, 0.5% octylglucoside (OCG) or no detergent, respectively. The virus from the four groups was purified using CsCl gradients and rAAAV genomes were quantitated using quantitative PCR. No effect was observed on yield of viral particles or density of rAAAV in the four preparations. After dialysis against PBS, transduction efficiency was measured by titration on CEF cells. Addition of OCG or CHAPS had no significant effect on transduction efficiency. However, deoxycholate which is a stronger ionic detergent reduced transduction efficiency almost 10-fold.

Tissue tropism of rAAAV was determined in CEF, DF1, LMH, DT-90, QNR, QT6, 293T, COS, primary chicken embryonic kidney cells, primary chicken pituitary cells and primary human fibroblasts and compared with that of rAAV2, rAAV4 and rAAV5 (Table 1). Table 1 shows the titers for rAAAV, rAAV2, rAAV4 and rAAV5 expressing LacZ in avian and mammalian cell lines and primary cells. Transductions were performed as described in Methods and Materials and efficiency was expressed as transducing units per  $10^6$  recombinant particles.

Transduction efficiency of rAAAV was 10-300 fold higher in avian cells compared with that of rAAV2, rAAV4 and rAAV5. In contrast, transduction of the mammalian cells in the panel by rAAAV was almost absent. This observation suggests that AAV is using a different uptake or transduction mechanism compared with the primate AAVs. Interestingly, rAAAV exhibited ~15-fold higher transduction efficiency in primary chicken embryonic fibroblasts compared to immortalized embryonic fibroblasts (Fig. 5B).

The present invention also showed that AAV ITR can function as a universal ITR for packaging with AAV2, 5 Rep proteins. Cross packaging experiments were carried out by transducing 293T cells with the two production plasmids (an ITR containing plasmid and a RepCap production plasmid) indicate and a third helper plasmid to supply adenovirus function. Forty-eight hours post transfection, cells were harvested and the amount of DNase resistance virus measured by quantitative PCR.

Table 1

Cell type	Transducing units per 10 <sup>6</sup> genomes			
	rAAAV	rAAV2	rAAV4	rAAV5
CEF	7140 ± 380	25 ± 3.5	84 ± 6.3	58 ± 5.7
DF-1	530 ± 35	8 ± 0.9	45 ± 4.7	60 ± 6.1
LMH	2380 ± 145	230 ± 25	34 ± 5.6	40 ± 4.9
DT-90	ND	ND	ND	ND
QNR	1260 ± 90	176 ± 18	42 ± 5.2	185 ± 26
QT6	930 ± 62	112 ± 21	23 ± 3.8	33 ± 5
Chicken Primary Embryonic Kidney cells	8080 ± 560	422 ± 46	350 ± 40	235 ± 38
Chicken Primary pituitary cells	4640 ± 375	144 ± 17	70 ± 12	91 ± 8.4
293T	ND	4500 ± 355	3130 ± 270	684 ± 57
COS	5 ± 0.7	6920 ± 420	3550 ± 165	592 ± 53
A549	ND	2190 ± 315	1360 ± 140	26 ± 4.3
Humary primary fibroblasts	ND	1990 ± 170	1130 ± 145	292 ± 31

5 Numbers represent the mean ± standard error from four independent transduction assays. ND=none detected.

#### Characterization of Binding and Transduction

10 The characterization of the binding and transduction requirements is important for the optimal utilization of a vector. Therefore we have examined the binding and transduction requirements of avian AAV (AAAV). To date, primate AAVs have been shown to require cell surface expression of either heparin sulfate proteoglycans (HSPG) (AAV2, 3) or sialic acid (AAV4, 5) for virus binding and attachment. However our

studies with avian AAV indicate that neither is required. Surprisingly, AAV required a distinct form of glycosylation, terminal lactose, for efficient binding and transduction, which is unique a distinct from that of the primate AAVs.

5 Initial experiments with AAV demonstrated that transduction is insensitive to competition with soluble heparin, which blocks binding with HSPG, soluble sialoconjugates, which blocks binding with sialic acid, or treatment with neuraminidase, which removes cell surface sialic acid (Figs 7, 8 and 6, respectively). Thus, Avian AAV appeared to be requiring a unique cell surface epitope. To  
10 characterize this epitope we treated DF-1 cells with several different inhibitors of glycosylation. Treatment with tunicamycin, which inhibits N-linked glycosylation, blocked both virus binding and transduction. In contrast, treatment with the O-linked inhibitor N-benzyl gal NAc had no effect (Figs 9 and 11, respectively).

15 Similar results were obtained with other N-linked inhibitors including N-butyl deoxynojirimycin and the unmodified form deoxynojirimycin. PDMP and Fumonisin B1, which inhibit the glycosylation of sphingolipids and ceramides, had no effect on AAV binding or transduction suggesting that the carbohydrate necessary for binding was attached to a protein (Fig 12). Lack of inhibition by Fumonisin B1 and PDMP and  
20 DGJ suggest lipids are not involved, but inhibition of binding and transduction with NB-DNJ and DNJ suggests glycoprotein is involved.

Initial experiments to block transduction by treatment with the protease trypsin had no effect on transduction (Fig 10). While trypsin is considered a broad specificity  
25 protease, its activity can be blocked by glycosylation; therefore we tested trypsin treatment after incubating the cells with low levels of tunicamycin which did not effectively block transduction. Treatment with trypsin or low levels of tunicamycin alone inhibited 0% or 25% of AAV transduction respectively. However, the combination of the two inhibited greater than 90% of AAV transduction confirming  
30 that AAV required the presence of a N-linked glycoprotein for efficient transduction

(Fig 10).

To further identify the carbohydrate component we tested a series of lectins for the ability to block virus binding and transduction (Fig 13). These lectins are briefly

5 described as follows:

MAA- This lectin binds glycoconjugates having galactosyl (b-1,4) N-acetylglucosamine structures. *Maackia amurensis* lectin I seems to tolerate substitution of N-acetylglucosamine with sialic acid at the 3 position of galactose;

10 SNA- *Sambucus nigra* lectin binds preferentially to sialic acid attached to terminal galactose in (a-2,6), and to a lesser degree, (a-2,3), linkage;

UEA-I- UEA I binds to many glycoproteins and glycolipids containing a-linked fucose residues;

15 PSA - This lectin has specificity toward a-linked mannose-containing oligosaccharides, with an N-acetylchitobiose-linked a-fucose residue included in the receptor sequence;

20 PHA-P - This lectin binds to complex carbohydrate structures on the cell surface;

MPL- This lectin prefers alpha linked N-acetylglucosamine structures;

25 EEL- This lectin has a carbohydrate binding specificity toward type 1 or type 2 chain blood group B structures but will bind other oligosaccharides containing galactosyl (a-1,3) galactose;

Con A- recognizes a commonly occurring sugar structure, a-linked mannose;

30 BPL- Binding appears to be highest for glycoconjugates containing galactosyl (b-1,3) N-acetylglucosamine structures but oligosaccharides with a terminal alpha linked N-acetylglucosamine can also bind;

35 ERCL - *Erythrina corallodendron* has an affinity for N-acetylglucosamine, N-acetyl-D-galactosamine, lactose and D-galactose;

40 WGA- The receptor sugar for WGA is N-acetylglucosamine, with preferential binding to dimers and trimers of this sugar. WGA can bind oligosaccharides containing terminal N-acetylglucosamine or chitobiose, structures which are common to many serum and membrane glycoproteins; and

WGA-s succinylated wheat germ agglutinin does not bind to sialic acid residues, unlike the native form, but retains its specificity toward N-acetylglucosamine (Eur. J.



Biochem. 98, 39, 1979 and Eur. J. Biochem. 104, 147, 1980).

In agreement with the neuraminidase data, lectins MAA and SNA, which bind sialic acid, had no effect on AAV binding or transduction (Fig 13). Furthermore, both WGA and the succinylated form, which does not bind sialic acid, both inhibited AAV binding and transduction in agreement with the MAA and SNA data. Binding and transduction were also inhibited by *Erythrina corralodendron* lectin which binds terminal poly lactose, suggesting that AAV may bind this carbohydrate complex. To test this hypothesis, competition experiments were carried out with soluble sialolactose conjugates or lactose complexes alone. While AAV was inhibited by the terminal lactose conjugates, AAV5 was not, confirming the results of the lectin blocking experiments (Fig 8).

Taken together, these results indicate that AAV requires N-linked terminal lactose present on cell surface proteins for efficient binding and entry. While other proteins may be involved in transduction, terminal lactose should be considered as a co-receptor for AAV binding and entry. This finding was completely unexpected and very different from that of primate AAVs.

Based on this data, the use of lactose affinity chromatography (e.g., columns) for the purification of AAV is provided. An example of lactose affinity chromatography is described by Tasumi et al., Primary structure and characteristics of a lectin from skin mucus of the Japanese eel *Anguilla japonica*, J Biol Chem. 2002 Jul 26;277(30):27305-11 (which is incorporated herein by reference).

Throughout this application, various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

## REFERENCES

1. Arella, M., S. Garzon, J. Bergeron, and P. Tijssen. Handbook of Parvoviruses. Vol. 1. ed. P. Tijssen. Boca Raton, Florida, CRC Press, 1990.
2. Bachmann, P.A., M.D. Hoggan, E. Kurstak, J.L. Melnick, H.G. Pereira, P. Tattersall, and C. Vago. 1979. *Intervirology* 11: 248-254.
3. Bantel-Schaal, U. and M. Stohr. 1992. *J. Virol.* 66: 773-779.
4. Chang, L.S., Y. Shi, and T. Shenk. 1989. *J. Virol.* 63: 3479-88.
5. Chejanovsky, N. and B.J. Carter. 1989. *Virology* 173: 120-128.
6. Chejanovsky, N. and B.J. Carter. 1989. *Virology* 171: 239-247.
7. Chiorini, J.A., S.M. Wiener, R.M. Kotin, R.A. Owens, SRM Kyöstiö, and B. Safer. 1994. *J. Virol.* 68: 7448-7457.
8. Chiorini, J.A., M.D. Weitzman, R.A. Owens, E. Urcelay, B. Safer, and R.M. Kotin. 1994. *J. Virol.* 68: 797-804.
9. Chiorini, J.A., C.M. Wendtner, E. Urcelay, B. Safer, M. Hallek, and R.M. Kotin. 1995. *Human Gene Therapy* 6: 1531-1541.
10. Chiorini, J.A., L. Yang, B. Safer, and R.M. Kotin. 1995. *J. Virol.* 69: 7334-7338.
11. Dixit, M., M.S. Webb, W.C. Smart, and S. Ohi. 1991. *Gene* 104: 253-7.
12. Fisher, R.E. and H.D. Mayor. 1991. *J Theor Biol* 149: 429-39.
13. Flotte, T.R., S.A. Afione, C. Conrad, S.A. McGrath, R. Solow, H. Oka, P.L. Zeitlin, W.B. Guggino, and B.J. Carter. 1993. *Proc. Natl. Acad. Sci.* 90: 10613-10617.
14. Flotte, T.R., S.A. Afione, R. Solow, M.L. Drumm, D. Markakis, W.B. Guggino, P.L. Zeitlin, and B.J. Carter. 1993. *J Biol Chem* 268: 3781-90.
15. Hermonat, P.L., M.A. Labow, R. Wright, K.I. Berns, and N. Muzyczka. 1984. *J. Virol.* 51: 329-339.
16. Hermonat, P.L. and N. Muzyczka. 1984. *Proc Natl Acad Sci USA* 81: 6466-70.
17. Hunter, L.A. and R.J. Samulski. 1992. *J. Virol.* 66: 317-24.

18. **Ito, M. and H.D. Mayor.** 1968. *J. Immuno.* **100**: 61-68.
19. **Janik, J.E., M.M. Huston, K. Cho, and J.A. Rose.** 1989. *Virology* **168**: 320-9.
20. **Kaplitt, M.G., P. Leone, R.J. Samulski, X. Xiao, D.W. Pfaff, K.L. O'Malley, and J.M. During.** 1994. *Nature Genetics* **8**: 148-154.
21. **Kotin, R.M., M. Siniscalco, R.J. Samulski, X. Zhu, L. Hunter, C.A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K.I. Berns.** 1990. *Proc. Natl. Acad. Sci. (USA)* **87**: 2211-2215.
22. **Laughlin, C.A., N. Jones, and B.J. Carter.** 1982. *J. Virol.* **41**: 868-76.
23. **Laughlin, C.A., M.W. Myers, D.L. Risin, B.J. Carter.** 1979. *Virology* **94**: 162-74.
24. **McCarty, D.M., J. Pereira, I. Zolotukhin, X. Zhou, J.H. Ryan, and N. Muzyczka.** 1994. *J. Virol.* **68**: 4988-4997.
25. **Mendelson, E., J.P. Trempe, and B.J. Carter.** 1986. *J. Virol.* **60**: 823-832.
26. **Mizukami, H., N.S. Young, and K.E. Brown.** 1996. *Virology* **217**: 124-130.
27. **Muster, C.J., Y.S. Lee, J.E. Newbold, and J. Leis.** 1980. *J. Virol.* **35**: 653-61.
28. **Muzyczka, N.** 1992. *Curr Top Microbiol Immunol* **158**: 97-129.
29. **Parks, W.P., J.L. Melnick, R. Rongey, and H.D. Mayor.** 1967. *J. Virol.* **1**: 171-180.
30. **Podsakoff, G., K.K. Jr Wong, and S. Chatterjee.** 1994. *J. Virol.* **68**: 5656-5666.
31. **Rose, J.A., M.D. Hoggan, F. Koczot, and A.J. Shatkin.** 1968. *J. Virol.* **2**: 999-1005.
32. **Russell, D.W., A.D. Miller, and I.E. Alexander.** 1994. *Proc. Natl. Acad. Sci. USA* **91**: 8915-8919.
33. **Ryan, J.H., S. Zolotukhin, and N. Muzyczka.** 1996. *J. Virol.* **70**: 1542-1553.
34. **Samulski, R.J., K.I. Berns, M. Tan, and N. Muzyczka.** 1982. *Proc Natl Acad Sci USA* **79**: 2077-81.
35. **Samulski, R.J., L.S. Chang, and T. Shenk.** 1989. *J. Virol.* **63**: 3822-8.

36. **Sanes, J.R., J.L.R. Rubenstein, and J.F. Nicocas.** 1986. *EMBO* **5**: 3133-3142.
37. **Senapathy, P., J.D. Tratschin, and B.J. Carter.** 1984. *J Mol Biol* **179**: 1-20.
38. **Tratschin, J.D., I.L. Miller, and B.J. Carter.** 1984. *J. Virol.* **51**: 611-619.
39. **Trempe, J.P. and B.J. Carter.** 1988. *J. Virol.* **62**: 68-74.
40. **Trempe, J.P., E. Mendelson, and B.J. Carter.** 1987. *Virology* **161**: 18-28.
41. **Walsh, C.E., J.M. Liu, X. Xiao, N.S. Young, A.W. Nienhuis, and R.J. Samulski.** 1992. *Proc Natl Acad Sci USA* **89**: 7257-61.
42. **Winocour, E., M.F. Callaham, and E. Huberman.** 1988. *Virology* **167**: 393-9.
43. **Jaksch, M., K.D. Gerbitz, and C. Kilger.** 1995. *Clin. Biochem.* **28**:503-509
44. **Burcin, M.M., O'Malley, B.W. and S.Y. Tsai.** 1998. *Frontiers in Bioscience* **3**:1-7.

It will be apparent to those skilled in the art that various modifications and variations can be made in the present invention without departing from the scope or spirit of the invention. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

What is claimed is:

1. A nucleic acid vector comprising a pair of avian adeno-associated virus (AAAV) inverted terminal repeats and a promoter between the inverted terminal repeats.
2. The vector of claim 1, wherein the promoter is an AAV promoter p5.
3. The vector of claim 1, wherein the p5 promoter is AAAV p5 promoter.
4. The vector of claim 1, further comprising an exogenous nucleic acid functionally linked to the promoter.
5. The vector of claim 1 encapsidated in an adeno-associated virus particle.
6. The particle of claim 5, wherein the particle is an AAAV particle.
7. The particle of claim 5, wherein the particle is an AAV1 particle, an AAV2 particle, an AAV3 particle, an AAV4 particle, an AAV5, an AAV6, an AAV7, or an AAV8 particle.
8. The particle of claim 5, wherein the particle is parvovirus particle.
9. The particle of claim 5, wherein the particle is dependent parvovirus particle.
10. The particle of claim 5, wherein the particle is parvovirus particle.
11. A recombinant AAAV virion containing a vector comprising a pair of AAAV inverted terminal repeats.

12. The virion of claim 8, wherein the vector further comprises an exogenous nucleic acid inserted between the inverted terminal repeats.
13. An isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:1.
14. An isolated nucleic acid consisting essentially of the nucleotide sequence set forth in SEQ ID NO:1.
15. An isolated nucleic acid that selectively hybridizes with the nucleic acid of claim 13.
16. An isolated nucleic acid encoding an AAV Rep protein.
17. The nucleic acid of claim 13, wherein the avian adeno-associated virus Rep protein has the amino acid sequence set forth in SEQ ID NO:3.
18. The nucleic acid of claim 13, wherein the AAV Rep protein has the amino acid sequence set forth in SEQ ID NO:5.
19. The nucleic acid of claim 13, wherein the AAV Rep protein has the amino acid sequence set forth in SEQ ID NO:7.
20. The nucleic acid of claim 13, wherein the AAV Rep protein has the amino acid sequence set forth in SEQ ID NO:9.
21. An isolated AAV Rep protein.

22. The isolated AAV Rep protein of claim 21, having the amino acid sequence set forth in SEQ ID NO:3, or a unique fragment thereof.
23. The isolated AAV Rep protein of claim 21, having the amino acid sequence set forth in SEQ ID NO:5, or a unique fragment thereof.
24. An isolated antibody that specifically binds the protein of claim 18.
25. An isolated AAV capsid protein.
26. The isolated AAV capsid protein of claim 25 having the amino acid sequence set forth in SEQ ID NO:11.
27. The isolated AAV capsid protein of claim 26 having the amino acid sequence set forth in SEQ ID NO:11 wherein SEQ ID NO: 11 has been modified.
28. The isolated AAV capsid protein of claim 27, wherein the modification alters the tropism of AAV.
29. The isolated AAV capsid protein of claim 27, wherein the modification results in a host immune response to SEQ ID NO: 11.
30. The isolated AAV capsid protein of claim 27, wherein the modification results in an immune response directed against AAV.
31. An isolated antibody that specifically binds the protein of claim 25.
32. The isolated AAV capsid protein of claim 25, having the amino acid sequence set forth in SEQ ID NO:13.

33. An isolated antibody that specifically binds the protein of claim 32.
34. The isolated AAV capsid protein of claim 25, having the amino acid sequence set forth in SEQ ID NO:15.
35. An isolated antibody that specifically binds the protein of claim 34.
36. An isolated nucleic acid encoding the protein of claim 25.
37. The nucleic acid of claim 36, having the nucleic acid sequence set forth in SEQ ID NO:10.
38. The nucleic acid of claim 36, having the nucleic acid sequence set forth in SEQ ID NO:12.
37. The nucleic acid of claim 36, having the nucleic acid sequence set forth in SEQ ID NO:14.
38. An isolated nucleic acid that selectively hybridizes with the nucleic acid of claim 36.
39. An AAV particle comprising a capsid protein consisting essentially of the amino acid sequence set forth in SEQ ID NO:11.
40. An isolated nucleic acid comprising an AAV p5 promoter.
41. A method of screening a cell for infectivity by AAV, comprising contacting the cell with AAV and detecting the presence of AAV in the cells.
42. A method of determining the suitability of an AAV vector for administration



to a subject, comprising contacting an antibody-containing sample from the subject with an antigenic fragment of a protein of claim 25 and detecting an antibody-antigen reaction in the sample, the presence of a neutralizing reaction indicating the AAV vector to be unsuitable for use in the subject.

43. A method of determining the presence in a subject of an AAV-specific antibody comprising, contacting an antibody-containing sample from the subject with an antigenic fragment of the protein of claim 25 and detecting an antibody-antigen  
5 reaction in the sample, the presence of a reaction indicating the presence of an AAV-specific antibody in the subject.

44. A method of delivering a nucleic acid to a cell, comprising administering to the cell an AAV particle containing a vector comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to the cell.

45. The method of claim 44, wherein the AAV inverted terminal repeats are AAV inverted terminal repeats.

46. A method of delivering a nucleic acid to a subject comprising administering to a cell from the subject an AAV particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, and returning the cell to the subject, thereby delivering the nucleic acid to the subject.

47. A method of delivering a nucleic acid to a cell in a subject comprising administering to the subject an AAV particle comprising the nucleic acid inserted between a pair of AAV inverted terminal repeats, thereby delivering the nucleic acid to a cell in the subject.

48. A method of delivering a nucleic acid to a cell in a subject having antibodies to AAV2 comprising administering to the subject an AAV particle comprising the

nucleic acid, thereby delivering the nucleic acid to a cell in the subject.

49. An isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO:20.

50. An isolated nucleic acid comprising the nucleotide sequence set forth in SEQ ID NO: 18.

51. A vector system for producing infectious virus particles having a characteristic of AAV comprising: at least one vector comprising a nucleic acid selected from the group consisting of a pair of AAV inverted terminal repeats, a nucleic acid encoding an AAV capsid protein, and a nucleic acid encoding an AAV rep protein.

52. The vector system of claim 51, comprising two vectors.

53. The vector system of claim 52, wherein the first vector comprises a nucleic acid encoding an AAV Rep protein and the second vector comprises a pair of AAV inverted terminal repeats.

54. The vector system of claim 52, wherein the first vector comprises a nucleic acid encoding an AAV capsid protein and a nucleic acid encoding an AAV Rep protein and the second vector comprises a pair of AAV inverted terminal repeats.

55. The vector system of claim 52, wherein the first vector comprises a nucleic acid encoding an AAV capsid protein and the second vector comprises a pair of AAV inverted terminal repeats.

56. The vector system of claim 55, wherein the second vector comprises a pair of AAV1 inverted terminal repeats.

57. The vector system of claim 55, wherein the second vector comprises a pair of AAV2 inverted terminal repeats.
58. The vector system of claim 55, wherein the second vector comprises a pair of AAV3 inverted terminal repeats.
59. The vector system of claim 55, wherein the second vector comprises a pair of AAV4 inverted terminal repeats.
60. The vector system of claim 55, wherein the second vector comprises a pair of AAV 5 inverted terminal repeats.
61. The vector system of claim 55, wherein the second vector comprises a pair of AAV6 inverted terminal repeats.
62. The vector system of claim 55, wherein the first vector further comprises a nucleic acid encoding an AAV1 Rep protein.
63. The vector system of claim 55, wherein the first vector further comprises a nucleic acid encoding an AAV2 Rep protein.
64. The vector system of claim 55, wherein the first vector further comprises a nucleic acid encoding an AAV3 Rep protein.
65. The vector system of claim 55, wherein the first vector further comprises a nucleic acid encoding an AAV4 Rep protein.
66. The vector system of claim 55, wherein the first vector further comprises a nucleic acid encoding an AAV Rep protein.

67. The vector system of claim 55, wherein the first vector further comprises a nucleic acid encoding an AAV6 Rep protein.
68. The vector system of claim 51, wherein the first vector comprises a nucleic acid encoding an AAV capsid protein and the second vector comprises a pair of AAV inverted terminal repeats.
69. The vector system of claim 68, wherein the first vector comprises a nucleic acid encoding an AAV1 capsid protein.
70. The vector system of claim 68, wherein the first vector comprises a nucleic acid encoding an AAV2 capsid protein.
71. The vector system of claim 68, wherein the first vector comprises a nucleic acid encoding an AAV3 capsid protein.
72. The vector system of claim 68, wherein the first vector comprises a nucleic acid encoding an AAV4 capsid protein.
73. The vector system of claim 68, wherein the first vector comprises a nucleic acid encoding an AAV5 capsid protein.
74. The vector system of claim 68, wherein the first vector comprises a nucleic acid encoding an AAV6 capsid protein.
75. The vector system of any of claims 68 to 74, wherein the first vector further comprises a nucleic acid encoding an AAV Rep protein.

76. The vector system of any of claims 51 to 74, wherein the second vector further comprises a promoter between the inverted terminal repeats.
77. The vector system of claim 76, wherein the promoter is functionally linked to an exogenous nucleic acid.
78. A vector comprising a pair of AAV inverted terminal repeats, a nucleic acid encoding an AAV capsid protein and a nucleic acid encoding an AAV Rep protein.
79. The vector of claim 78, encapsidated in an dependent parvovirus particle.
80. The vector of claim 79, wherein the particle which encapsidates the vector is an AAV1 particle, an AAV2 particle, an AAV3 particle, an AAV4 particle, an AAV5 particle an AAV6 particle, and AAV7 particle, and an AAV8 particle.
81. A cell comprising an AAV nucleic acid.
82. A population of cells comprising an AAV nucleic acid.
83. A method of producing AAV virus particles comprising a) transducing a cell with the nucleic acid encoding AAV and b) isolating AAV virus particles.
84. The virus particle produced by the method of claim 83.
85. A cell comprising the vector system of any of claims 51-77.
86. A method of producing AAV virus particles comprising a) transducing a cell with the vector system of any of claims 51-77 and b) isolating AAV virus particles.

87. The virus particle produced by the method of claim 86.
88. A method of inducing an immune response to AAV in a subject comprising administering an AAV particle comprising SEQ ID NO: 11, wherein SEQ ID NO: 11 comprises an epitope that induces an immune response in a subject.
89. A method of blocking an immune response against AAV in a subject comprising administering an AAV particle comprising SEQ ID NO: 11, wherein SEQ ID NO: 11 comprises an epitope that blocks the immune response to AAV in a subject.
90. A method of producing a recombinant protein, comprising administering an AAV particle comprising an exogenous nucleic acid encoding a protein to an embryonated avian egg; and b) purifying the protein from the egg.
91. The protein purified by the method of claim 90.

1/22



FIG.1A

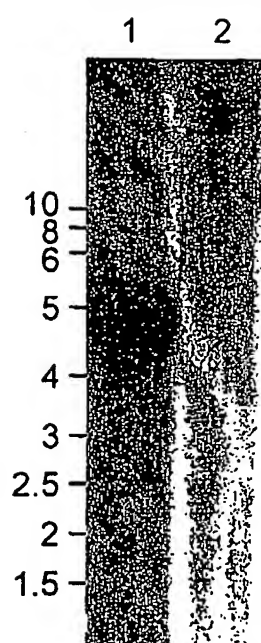


FIG.1B

2/22

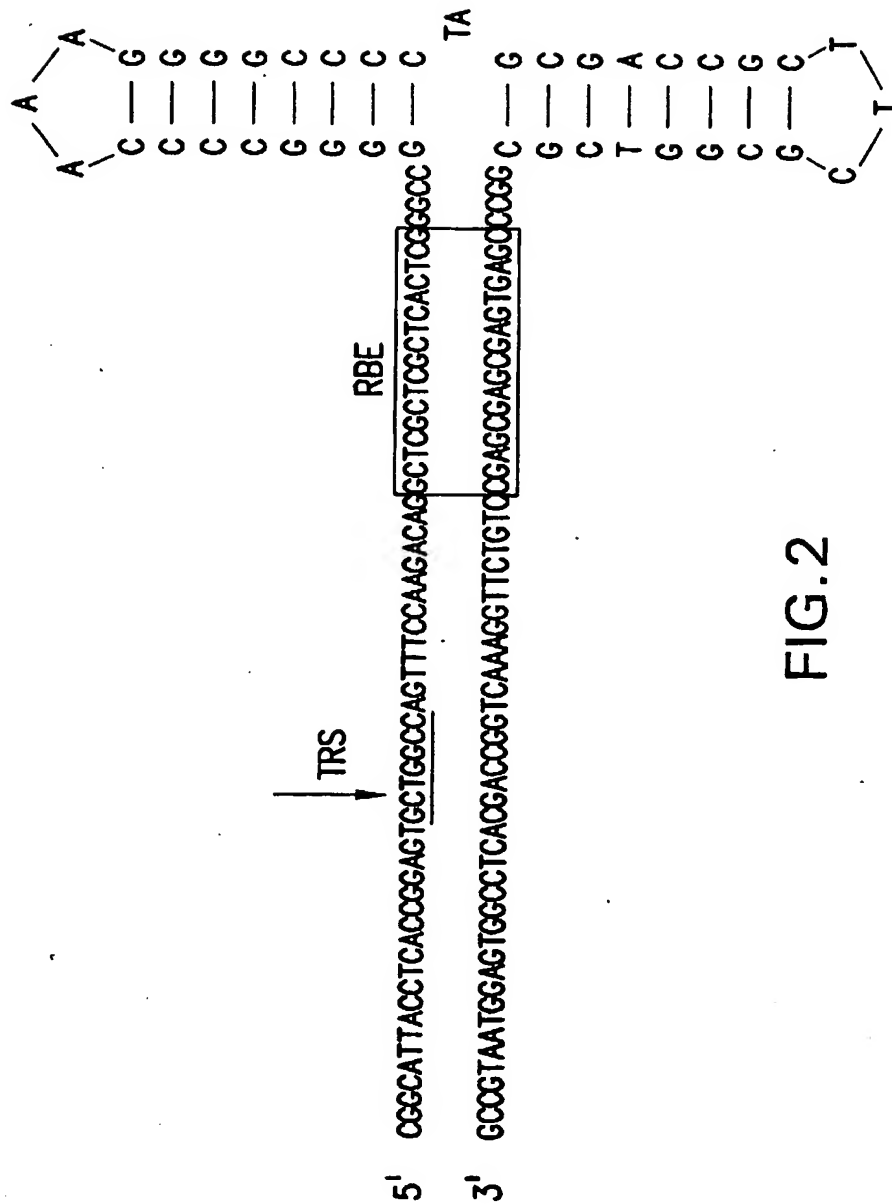


FIG.2



[illegible]

**FIG. 3A**

[illegible]

**FIG. 3B**

5/22

[illegible]

**FIG. 3C**









9/22

AAV1 GCAAAACACCGCTGGCTTTAGCAGACCGTCTACGATCAAGGACCGCCACGACCGGATCGTAACAGATCTCATCCAAAGAGAGCAAAATCAGACCCACCAACTCGGTGGTATCGA 3994  
 AAV2 -AGCGGGGTCTCATCTTTGGGAAGCA---AGGTCAGAGAAAACAATGTGAACATTGAAAGGTGATGATT---ACAGACGAGAGAGAAATCGGAACAACCAATCCCGTGGGTACGGA 3918  
 AAV4 CAACAGCCAGCTCATCTTTGGGGGGCT--AAACAGAAAGGCAACAG-GCCACCGTACCCGGGACTCTGATCTTCACTCTTGAGGAGGAGCTGGCAGGCCACCAAGCCACCGATACGGA 3977  
 AAV5 GAACACTATGATCTTCAACAGCCAGCGGGGAACCCGGGCAACACAGCCACGACCTCGAGAGGCAACATGCTCATCACCAGCGAGGAGAGAGCGAGCCGGTGAAACCGCGTGGGCTACAA 3902  
 \* \* \* \* \*  
 AAV1 CGGTGGGGAGGAGTTCCCAACCAACACAGTCAGTGTGACCCCGGGCACTCGCGGGCCGTCAACAATCAAGGGGGCGTTCCCGGGATGGTGTGGGCAAAACAGAGACATTTACCCCTAC 4114  
 AAV2 GCAGTATGTTCTGTATCTACCAACTCCAGAGAGGCAACAGACAGAGAGTGTCAACAGACAGAGGTTCTCCAGGATGGTCTGGCAGCAGAGAGATGTGTACCTT-C 4037  
 AAV4 CATGTGGGCAACCTTACCTTGGGGGTGACAGAGCAACAGCAACCTGCGACCGTGGAGAGCTGAGACGCTTGGAGCCGTGCTGGAAATGGTCTGGCAAAACAGAGACATTTACTAC-C 4096  
 AAV5 CGTGGGGGGAGATGGCCACCAACACAGAGCTCCACCACTGCCCCCGGAGCGGACGCTACAACCTCCAGGAAATGTTGGCCCGGAGCGGTGGATGGAGAGGGAGCGTGTACCTC-C 4021  
 \*\* \* \* \* \* \*  
 AAV1 AGGACCCCATTTGG-CCAAATTCGCGACACTGACATCACTTCCATCCGTCCCGGCTTATTTGGGGGGTTTGGCTGCAAGCATCCCGCTCCCGAGATTTTCAATAACACACACCCGTCC 4233  
 AAV2 AGGGGCCCATCTGGGCAAGATTCCACACAGCGGACGACATTTTCACCCCTCTCCCTCATGGGTGSAATTCGSACTTAAACACCCCTCTCCACAGATTTCTCATCAAGAACACCCCGGTAC 4157  
 AAV4 AGGTTCCATTTGGGGCAAGATTCTCATACGATGGACACTTTTCAACCTCACCGCTGATTTGGTGGTTTGGGCTGAAACACACCGGCTCTCAAAATTTTATCAAGAACACCCCGGTAC 4216  
 AAV5 AAGGACCATCTGGGGCAAGATCCAGAGACGGGGGGGCGCACTTTCAACCCCTCTCCGGCATGGGCGGATTTGGGACTCAAAACACCCACCGCCATGATGCTCATCAAGAACACCGCTGTGC 4141  
 \* \* \* \* \*  
 AAV1 CTGCCAACCTTCGGAAGGTTCCAGAGGCGCAAGTGGGCTCTCTCATCAACAGTACTGSAACCGGACA-GTGACCGTGGAAATCTTTGGGAACTCAAGAGSAAACCTTCCAAGCGC 4352  
 AAV2 CTGCGAATCCTTCGACCACTTCAGTGCGECAAAGTTTGCTTCTTCATCACAGTACTCCAGGAGACGCTCAGCGTGGAGATCGAGTGGGAGTGCAGAGSAAACAGCAACGCG 4277  
 AAV4 CTGCGAATCCTTCGACCACTTCAGTCTACTCCGTTAACTCTTCTTACTCAGTACAGCACTGGGCA-GGTGTGGTGGAGATTGACTGGGAGATCCAGAGGAGCGGTCCAAAGCG 4335  
 AAV5 CCGGAAATATC---ACCAGCTTCTGGAGGTGGCGGTCAGCAGCTTCATCAACAGTACAGCAACCGGCA-GGTCAACCGTGGAGATGGGAGTGGGAGTCAAGAGSAAACCTTCCAAGAGG 4257  
 \* \* \* \* \*  
 AAV1 TGGAACCCCGAAATCCAGTTACCTCCAACTTT---GGCAAGGCGGCGA-CATCCAGTTTGGGCTTCGACAGGSAATCCTATTCCGAACCTGTGCCATCGGTACCGGTTACCTTA 4467  
 AAV2 TGGAAATCCCGAAATTCAGTACACTTCCAACTAC---AACAGTCTGTTAATGCTGGACTTA-CGGTGGATAGTAAATGGGCTGTTATTCAGAGCTCGCCCATTTGECACAGATACCTGA 4392  
 AAV4 TGGAACCCCGAGGTCCAGTTTACCTCCAACTACGAGCAGCAAAACTCTCTGTTGGGCTCC-C---GATGCGCTGGGAAATACACTGAGCTTAGGGCTATCGGTACCGGCTACCTCA 4450  
 AAV5 TGGAACCCAGAGATCCAGTACACAAACAACTAC---AAGCA-CCCCAGTTTGTGGACTTTGGCCCGGACAGCAACCGGSAATACAGAACCCAGACCTATCGGAACCCGATACCTTA 4372  
 \*\*\*\*\*  
 AAV1 CCAAACCTCTGTAA-----ATTAA-----ACCTTCAA-TAAACG---TTATGGGTAACTGTATTTCCGTCTC---CTGTGTTTATTCAGTCACATGA----- 4550  
 AAV2 CTGTAATCTGTAAATGCTGTGTTAATCAATAAACCGTTTAAATGCTTTCAGTTGSACTTGGTCTCTCGGTTATTTCT-TCTATCTAGTTTCCATGGCTAG---GT-AGATAAGTAGC- 4506  
 AAV4 CCCACCACCTGTAAATACCTGTTAATCAATAAACCGGTTTATCGTTTCAGTTGSACTTGGTCTCGGTTGCTCTTATCTTATCTGTTTCCATGGCTACTGGGT-ACATAAGCAGCG 4569  
 AAV5 CCGGACCCCTTTAA---CCCATTCAT-----GTGGCATACCTCAA--TAAACGCTG-TATTGGTGTGAGTAAATACCTGCTC-----TTGTGGTCAITTCAT-GAATAA-CAGC- 4470  
 \*\* \* \* \* \* \*

Poly-A

VP-stop

FIG.3G





11/22

AAAV	-----MRSYVEIVQLPNDVESQVPGISDSFVNWITSREMTLPEDADMDLQDVQVQLTLGDKIQREIRTHGTMKPEPDFHYFIOLEGEVFFHLHLLTCSVKPMLGRYIR	110
AAV2	-----PGF IVIKV S LDGHL	110
AAV4	-----PGF IVIKV S LDGHL	110
AAV5	-----ATF RV F EHL	109
GP	-----MALSRPLQISSDKF IR SS IDQD L LN E LSTGV EPTG I NMEH NLPW AE KNIFQR NOFNQD ETDF F E SEYI CCTAQGN RSF MS	117
AAAV	-----HIQKIVSKVYCATSLRUKDGCWTKTN FGGANKVRAESYIPAYLIPKQOPEVQWNTWPEYIKACHREL RASLARLHFEAGVSKENLARTADG APWMPTRVSKRYMELVDW	228
AAV2	-----Q RE LIQRI RGIETPLNWEA R GA G VD C N L T L MEQ LS NLTE KR VAQ LTHVSQT EQNKENQPNSD IRSKT A	230
AAV4	-----Q KE L TRI RGVPEQLNWEA R GA G VDDC N L T L MDQ S NLAE KR VAQ LTHVSQT EQNKENQPNSD IRSKT A	230
AAV5	-----Q RAQL KV FQIEPQIN WVAI V K VDSG L V L LD KL A NL E KR VAQFLA SSQRS Q AASQREFSAD IRSKT OK A N	226
GP	-----Q KDS IRD EKGQIKIP WFAI R Q TVTAA LH K L F M LFTA A CLOK QE LDAFQ SOLAAP LPDPQ STV LISN AA N SN	232
AAAV	-----LVEKGITTEKEMLENRESFRSQASSNSARQIKTALQGAIQEMLLTKTAEDYLVGKDPVSDDDIRQNRITYKILELNHYDPAYVGSILVGHQCKKAGKRNLTMLFGHATTGKTNIAEIA	348
AAV2	-----D S Q IQ DOA YI N A RS A DN GKI S P QQ E SS G Q AA VFL AT F I P	348
AAV4	-----DR S Q IQ DOA YI N A RS A DN SKI S P QN P E SS R M G Q AA VFL A F I P	348
AAV5	-----H S Q IQ Q YL NSTG RS A DN TKI S S V SSVP E SK WQ F M G A Y RSFN V Y P	344
GP	-----I M S Q T Y T SNV V A EN RA T I L TK V Q KM N N Q I C VKREFN AI Y P	350
AAAV	-----HAPFYGCYNWNTNENFPNDCEKMTIWEEGKTAKWVETAKAILGGSRRVRDQCKASVIPETPTVITSNTMCVWIDGNTTTFEKPLEDKLELLTRLPDFGKVTQOEVRQ	468
AAV2	-----T D V S K S AQ D V A S Q Q Q F TR OH KD	468
AAV4	-----D V S K S AQ D V A S Q Q Q F TK EH KD	468
AAV5	-----T D L N S S Q DS V V V S Q F TK P I KD	464
GP	-----D L N S A G C D MIV S M RI E QIV SHK EPS IS K E 470	
AAAV	-----FFRMSQDLTPVITPEFLVRKAESRKRP APSGEGYISPTKRPALAEQQQASESADPVP TRYRIKCSKHGQDKMLFPQOICESMNRDINICAIHKTDDCK	567
AAV2	-----AK VVE EH Y K GGAK DADISE KRVRESVA PST DAEGASIN YAD QN R V NL RQ R QNS FT GOK L	569
AAV4	-----AS V E TH Y GGA NDADISE KRACPSVA PST DAEG D YAD QN R V NL RQ R QNVO FT GVM A	569
AAV5	-----A AKVNOV TH K PRELAGTYGAESLKRPLGDVNTSYKSLKRAKLSFVPETPRS DWT DPAPLRPLMNS DC DY AQF NTSNK DE YL GK G IC NV H Q	584
GP	-----K AN N V VS K TN QTNLP E VP RANE EEP KTWAPPTRE LEELLR ASPELFSSVAPITVTPQNSPE KRSRNNYQVRCALHTYDINSM VF	571
AAAV	-----ECFPDYGKDDVLPCTEHNVSRCYCHSGEL YRVTSDSKAPESDEGTEPSVAPCTIHLMGKSHGLVTCACRLKNSTLHDDLQGGDLEQ	662
AAV2	-----VS V E Q VSWKK AVOKL Y I VP D A T D V V CIF 621	
AAV4	-----VS V E Q VSWRKRTYQKL P I RAP E A S E A V CDM 623	
AAV5	-----I HGIP IPPWE K N SDFG F ANK 610	
GP	-----MECEKAN FP FQPLG N COEHGNYDCAICKELKNEL EI EHWFE AEN 627	

FIG.4A

12/22

AAV1	MSLISDIPDLERL VKKGVNAADFVHLESGPRPKANQDQTO-----ESLEKDSRGL VFPQYNYLGPFGNGLDKGEPVNEADAAALEHDKAVDLEIKDGHNPYFEYNEADRRFERLKL	114
AAV2	A-ADGYL.....DTLSEGIQAMK.KP.P.PAERHK.....L.K.....ROLDS.D.LK.H.AE.....Q	105
AAV4	T-DGYL.....DNLSEGVREWA.QP.A.K.H.....NA.L.K.....QQL.A.D.LK.H.AE.Q.Q	104
AAV5	-FVHP.....EVEGLE.LG.A.K.P.H.....OA.L.....R.EV.R.IS.NEQLA.D.LK.H.AE.K.A	104
GP	TFLSFEY.Y.....T.SMRN.KA.A.Q.P.S.SVSPDR.PER.NN.F.L.K.G.P.K.SV.....QQL.A.D.TKF.H.QD.IDS.Q	113
VP2		
AAV1	DOTSFGNGLGATFOAKKRVLPEFGLVEDS-KTAPTDGKRGKGEDEPLPDTSSQTPKPKKPKRPERPSGGAEDPGEGTSSNAGAAAPASSVG--SSINAEAGGGGPGVGDAGQAGVGNS	231
AAV2	E.....R.V.....EPV.....GKKRPVSESPVDPSSGTGKAGQAPARKLNFEQTG.ADSVPDQPLQG.APSGLGNT.T.S.A.MA.NNE.....T.S.A.MA.NNE	225
AAV4	G.....R.V.....QAGE.....GKKRPLI.SPQPDSSGTGKAGQAPARKLNFEQTG.ADSVPDQPLQG.APSGLGNT.T.S.A.MA.NNE.....T.S.A.MA.NNE	219
AAV5	V.....V.....EGA.....KRIDJHFK.KKARTEEDS.PSTSSDA.AGPSSQQLQIPAPASSLQATM.A.....L.NN.....A	215
GP	Q.....Q.....PWN.....AKKMTGKLTDPVW.....KLTEVSAGGSSAVQDGG.T.EGTEPIA--A.E.....AM.SSG.....A	221
VP3		
AAV1	GMHCSQMLENGWTRITRTWLPSPNNHLYKRIQGPS-GEONNNKFFGFSTPWGFDYNNRFCHFSRDMQRLINNNGIRPKAVRFLFNIOVKETVODFNTTIGNULTSTVQVFA	350
AAV2	T.MEDR.I.TS.....A.T.....Q.SSQ.....AS.D.HY.Y.....F.....RLN.K.....QN.GI.A.....T	344
AAV4	D.....T.S.GH.T.TS.....T.....LGE.....SLOS.TYN.....F.....M.....VKI.....TSNGE.VA.....I	335
AAV5	D.....T.MEDR.KS.....Q.RE.KSG.VD.S.A.AY.Y.....F.S.W.....Y.F.RSL.VKI.....ST.A.....T	335
GP	MG.T.I.K.....I.A.TSGT.SQ.A.VQYA.Y.....F.....H.....SLK.KI.V.....T.QIK.A.....I	340
AAV1	DKDYQLPYVLGSATEGTFPPADITYTIPQVGYCTLNNYNEAV--DRSAFYCLDYFSPDMRLTGNINFETTYTFEDVPFHSFMAHNOTLDRMLNPLVDYLVAFSSVSOA---GSSGRAL	464
AAV2	SE.....HQ.CL.....VFW.....LTLNGSOAVG.S.E.Q.....T.S.....SY.S.S.....I.....Y.L.RTNPSGTTQS-R	460
AAV4	SS.E.....MDAGQ.SL.....N.VFW.....G.VTG.TSQQT.N.....E.Q.....I.S.K.....Y.S.S.....I.....GLQ.TTTGTLNAGTATT	455
AAV5	D.....V.NG.....CL.A.PQVF.L.....ATLNDRONTENPTE.S.F.E.K.....N.E.....S.PS.N.FK.A.....YR.V.TNN---TGGV	447
GP	DEH.....M.....S.V.AL.....MHT.QNGAREN.....E.Q.....FD.E.....S.D.....N.NE.....D.S.NA	453
AAV1	HYSRATKTNMAOYRNWLPGEFRDQDQIFTGASNTKNVFSVME-KGQWELDRTNLMDGPAAATTSGEPRDQAMONTLAFSRIVDQDTATDRNQILITNEDEIRPTNSVGIDA	578
AAV2	OF.U.GASDTRD.S.....CY.Q.....RVSKTSADNNSEYSGTGA-----TKVHNGRSLVN.GP.MASHKDEEKFPOSQVILIFGKQSGKTNV.TEKVM.D.E.....T.P.ATEQ	570
AAV4	NFTKLRP.FSNFK.....SIKQ.GF.SKT.QNRY.IPATGSDSLIKYETHSTLQGRASALT.GPMMATA.PA.SKFNSQ.I.AGPKQNG.....VP.GTLIF.S.E.LAA.ATDT.M	569
AAV5	OFKNLAGRY.NT.K.F.....MG.T.GHNL.SGNIRASVSFAFATT--NRMELEGASVQVPPQPNQNTNLO.SNTYALENTMIFN.OPANPG.....YLEGNM.....S.S.TQ.V.R.AYVW	560
GP	DFKK.V.GAVGTMG.....K.L.....RVIRAYTGTGTDNYANWNI.S.N.NKWN.KD.QY.L.....VS.....YTE.....ASSLPA.....I.GIAKDP.RSGST.AGISD.MV.E.Q.VA.....G.WKP	567
Heparin binding		
AAV1	HEAVPTNNOSIVTPTGTRAAVNNQSGALPGMMWQNRDIYPTGTGLAKIPDQTNHFHPSPLIGRFGCKHPPQIFIKNTPVPANPSETFQTAQVASFINOYSTGQCTVIEIFMELKKETSKRW	698
AAV2	Y.S.S.L.RGNRDAAT.D.T.V.....D.V.LQ.PIW.....H.G.....M.G.L.....L.....T.SA.F.....T.....VS.....E.Q.N	690
AAV4	NL.GGD.NSNLP.VDRLTAL.....YQ.PIW.....H.G.....G.L.....G.....AT.SSTP.N.....T.....VS.Q.D.IQ.R.....	689
AAV5	G.QMA.....ST.APATGY.L.EIV.....S.ME.V.LQ.PIW.....E.GA.....AM.G.L.....MML.....G.-ITS.SDVP.S.....T.....V.....ME.....N	679
GP	Y.RTV.E.NTT.AP.SSDOLV.....LQ.PIG.....K.GK.....NL.G.LHN.....V.....D.PVEYVHQ.MN.Y.T.....MV.....R.N	687
AAV1	PEIQTSNFGWAQIOFANSDTGSYSEPRPIGTRYLTKPL	743
AAV2	Y.....Y.NKSNVND.T.DTN.V.....RN.....	735
AAV4	V.....Y.QQNSLLW.PDAA.K.T.A.....HH.....	734
AAV5	Y.N.YNDPQFVD.POS.E.RTT.....R.....	724
GP	S.RTS.M..PNE..G.V.D.L.....QN.....	732

FIG.4B

13/22

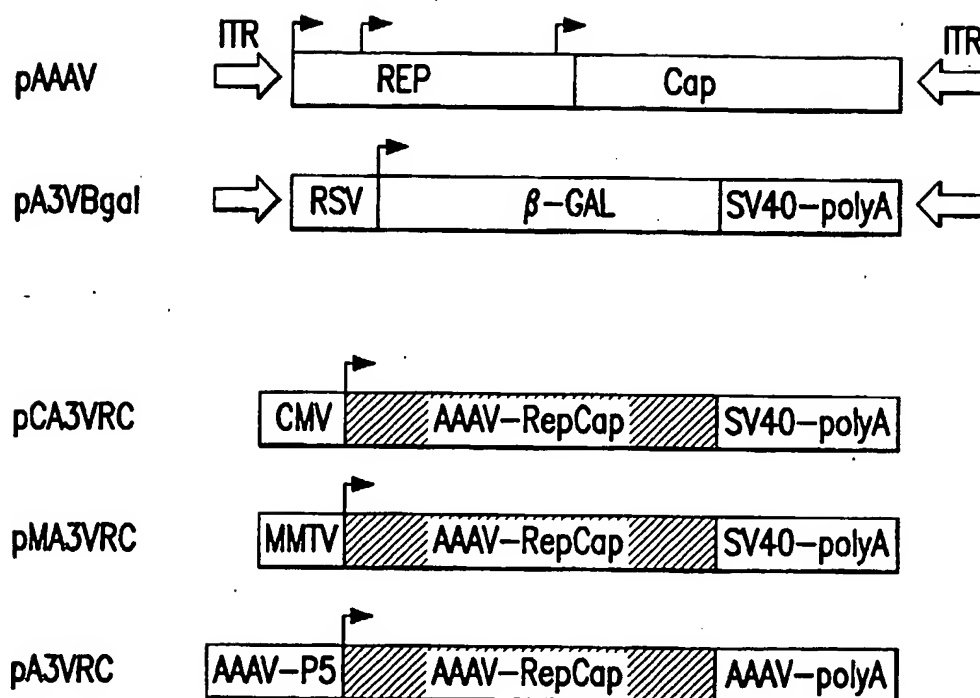


FIG.5A

14/22

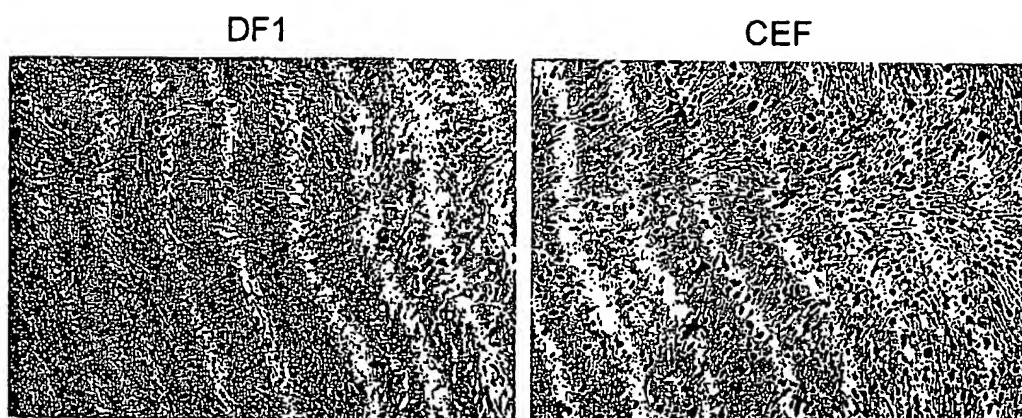


FIG.5B

15/22

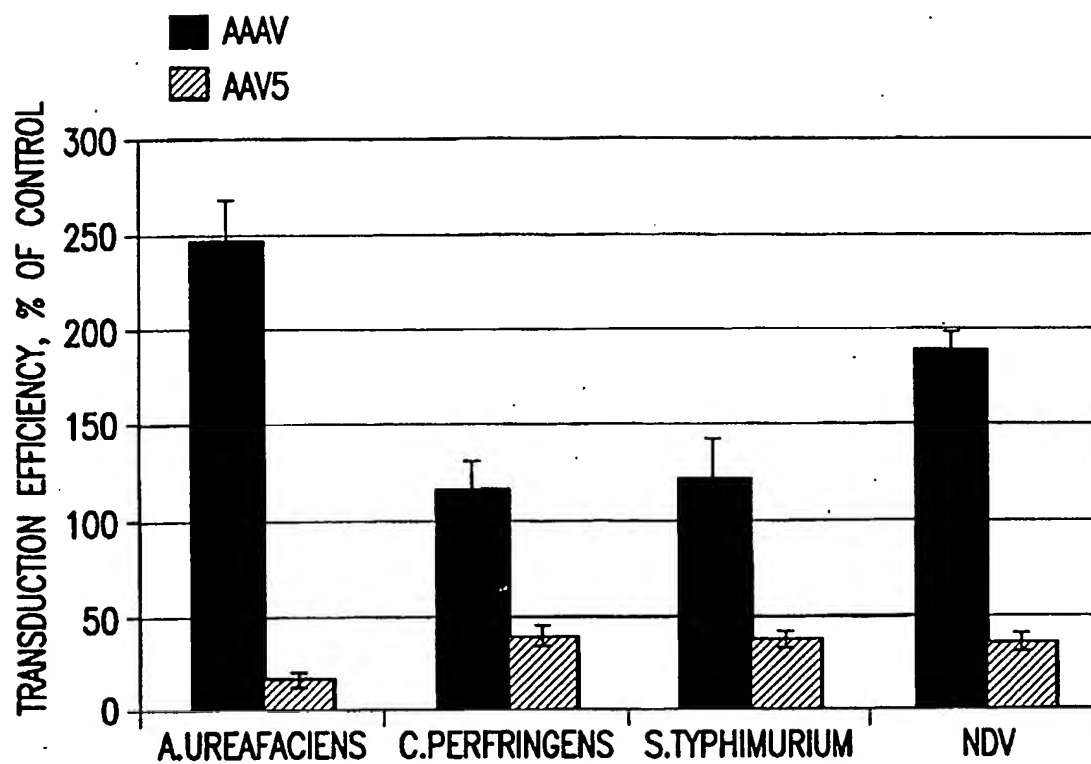


FIG.6

16/22

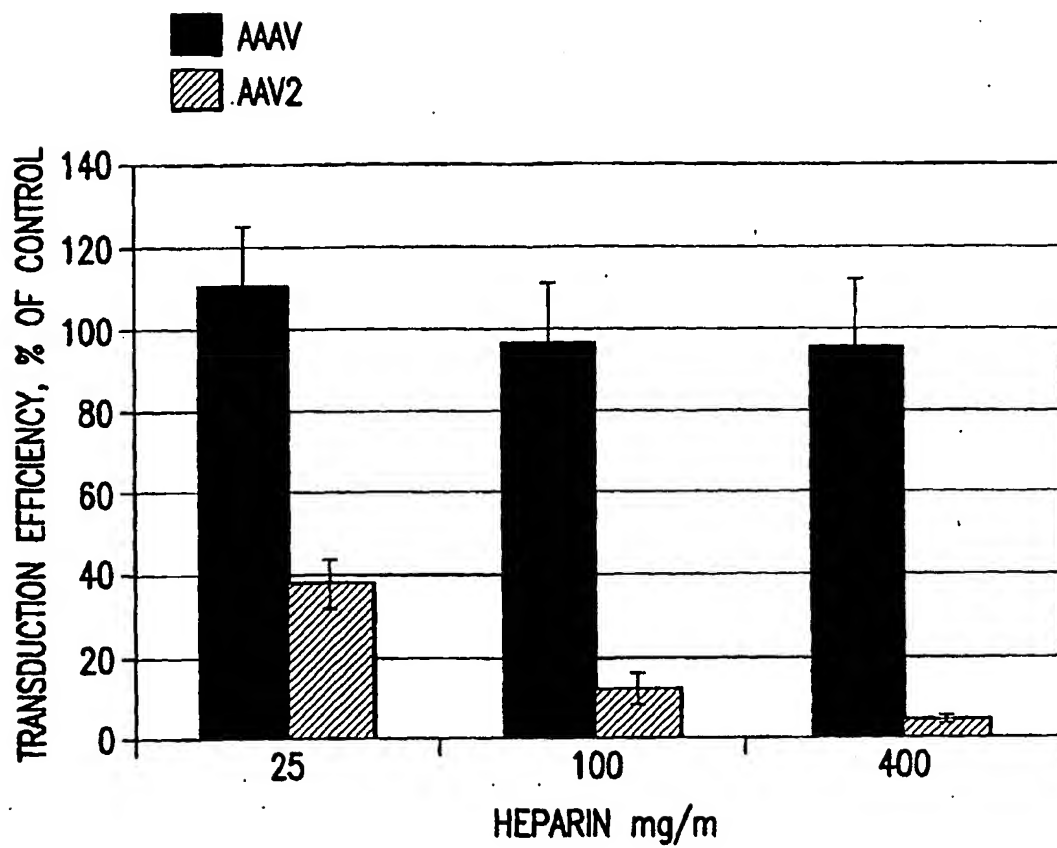


FIG.7

17/22

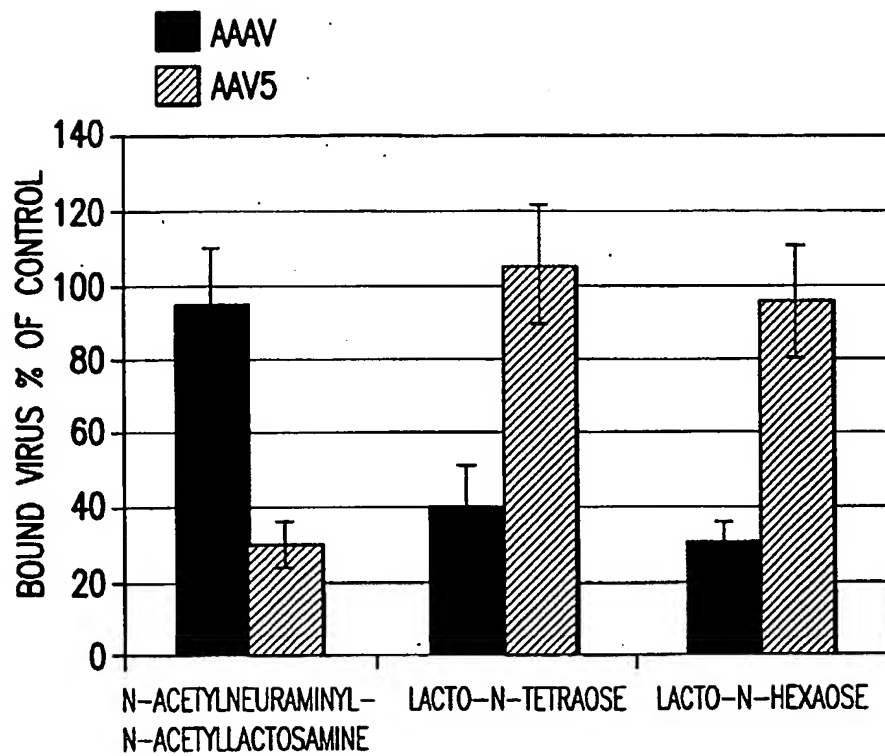


FIG.8

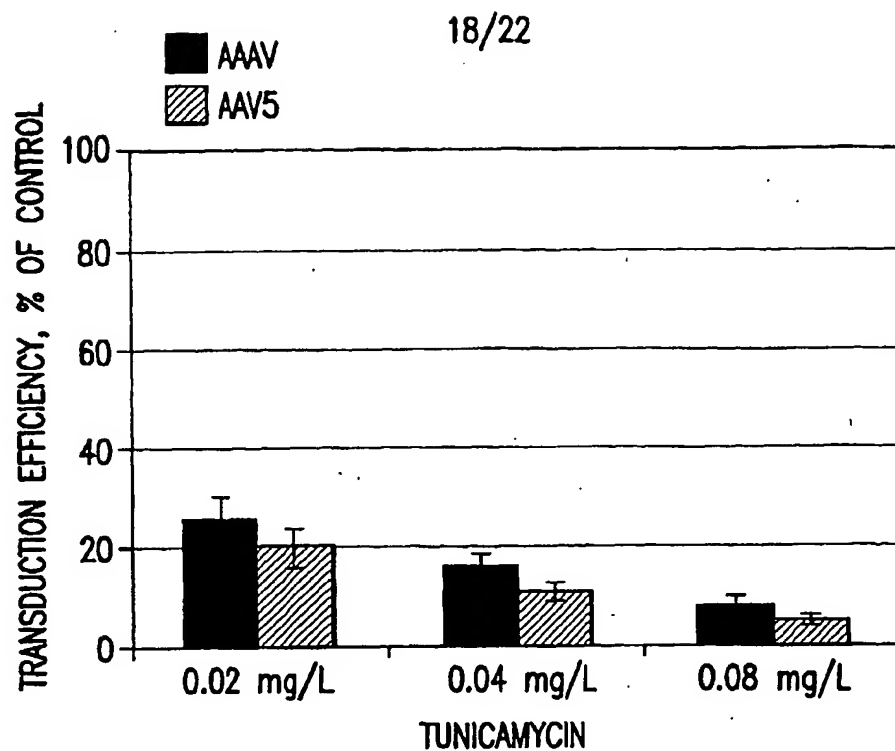


FIG.9A

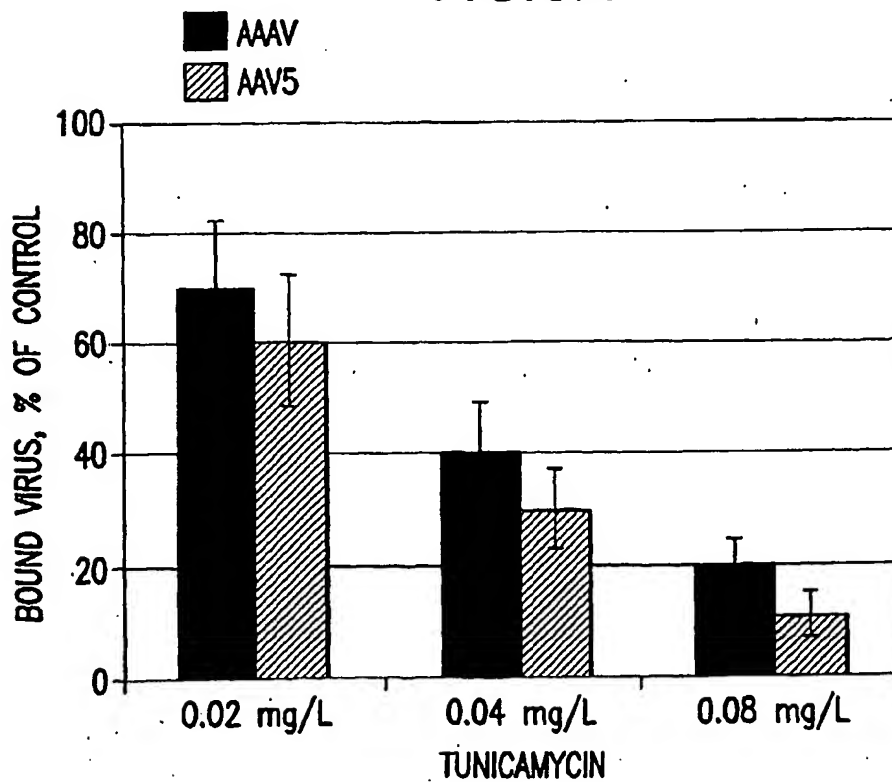


FIG.9B.



19/22

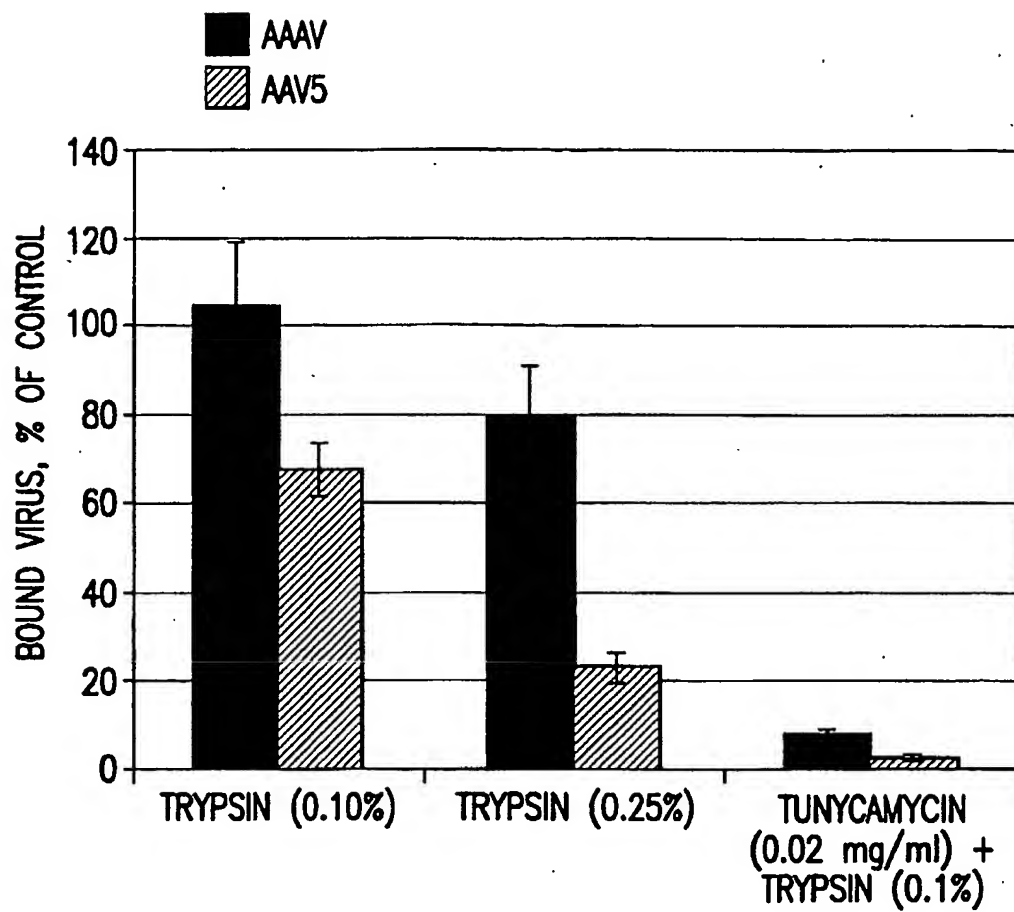


FIG. 10

20/22

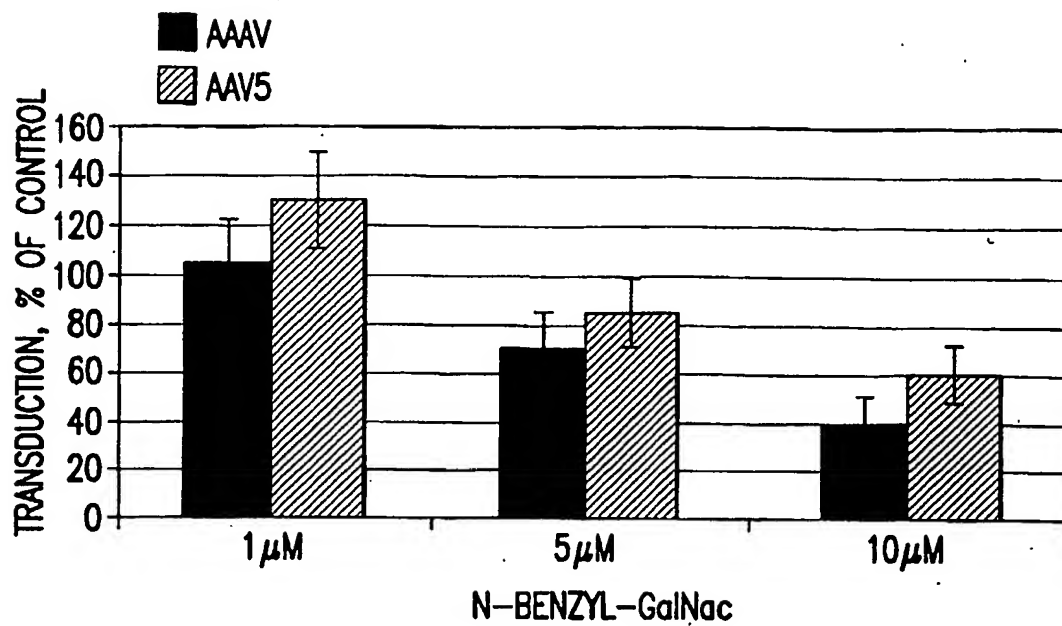


FIG. 11A

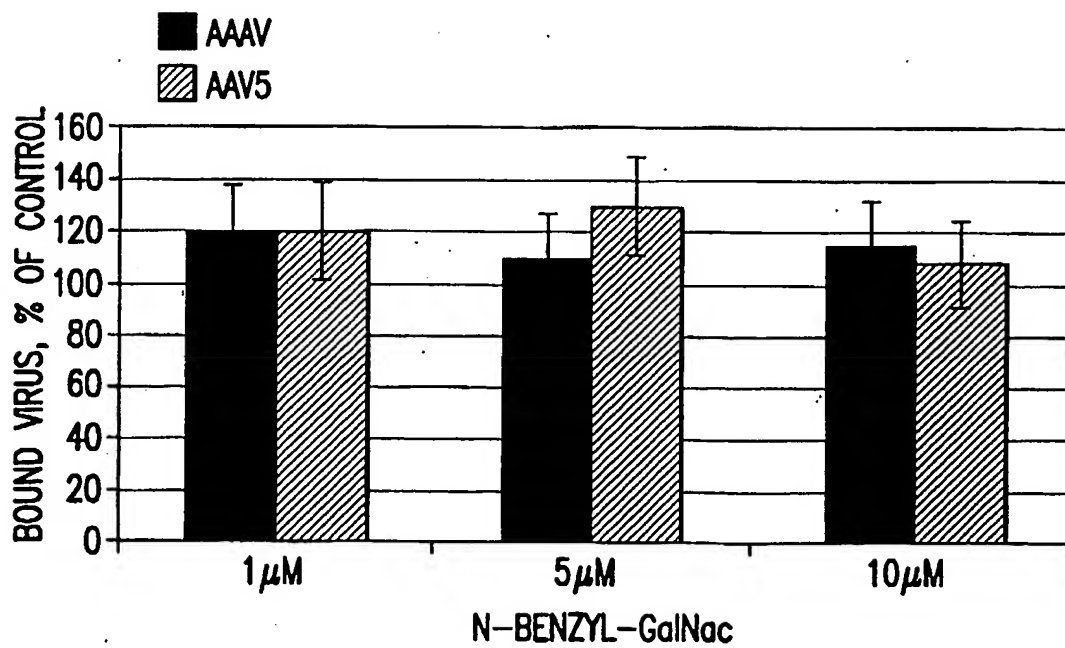


FIG. 11B

21/22

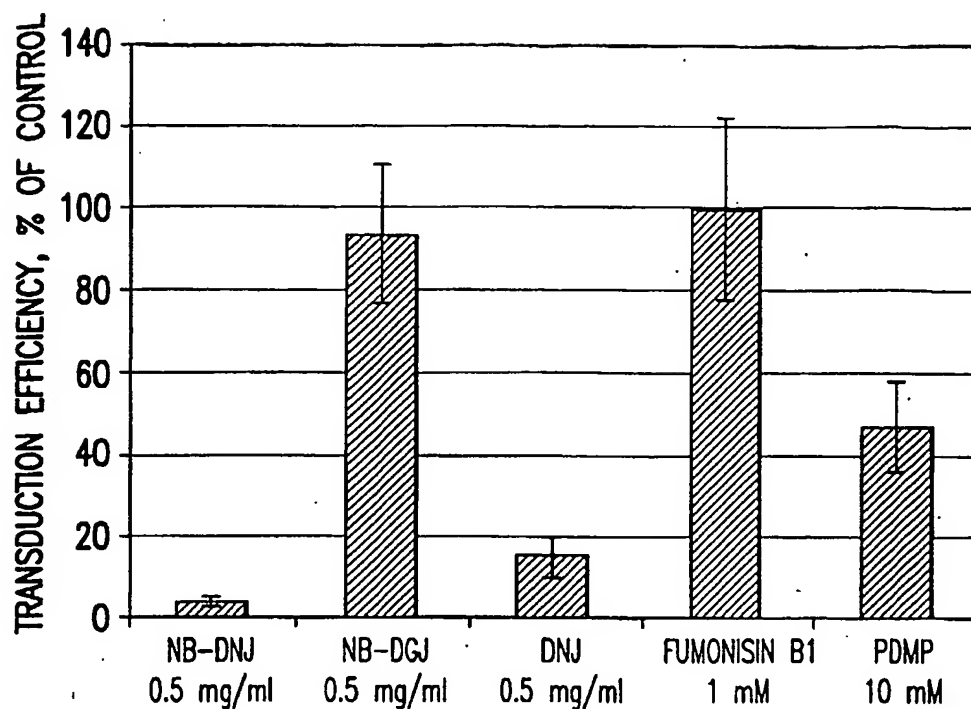


FIG.12A

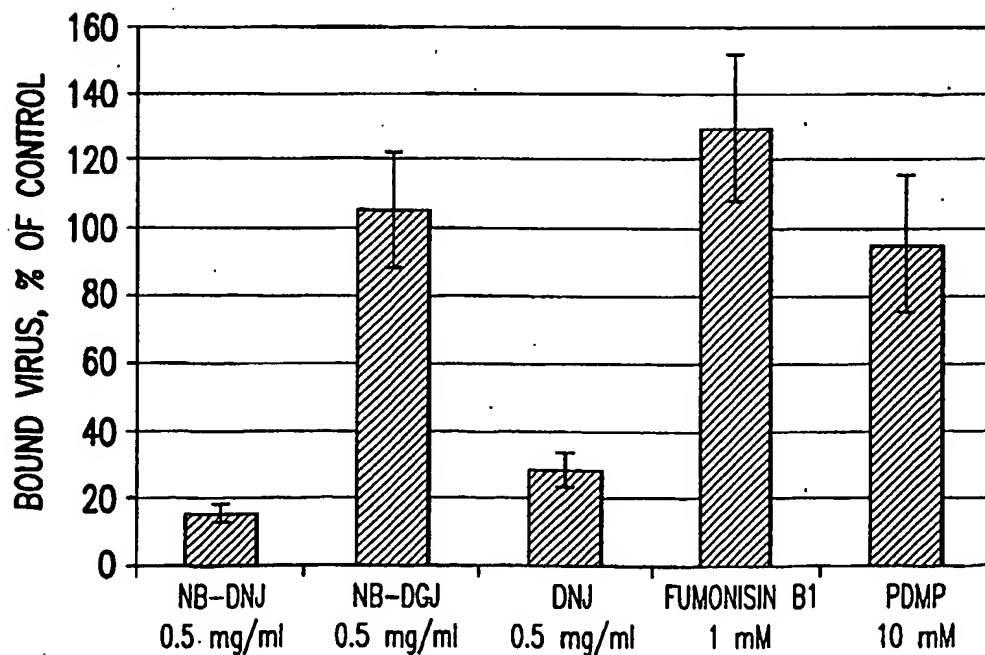


FIG.12B

SUBSTITUTE SHEET (RULE 26)

22/22

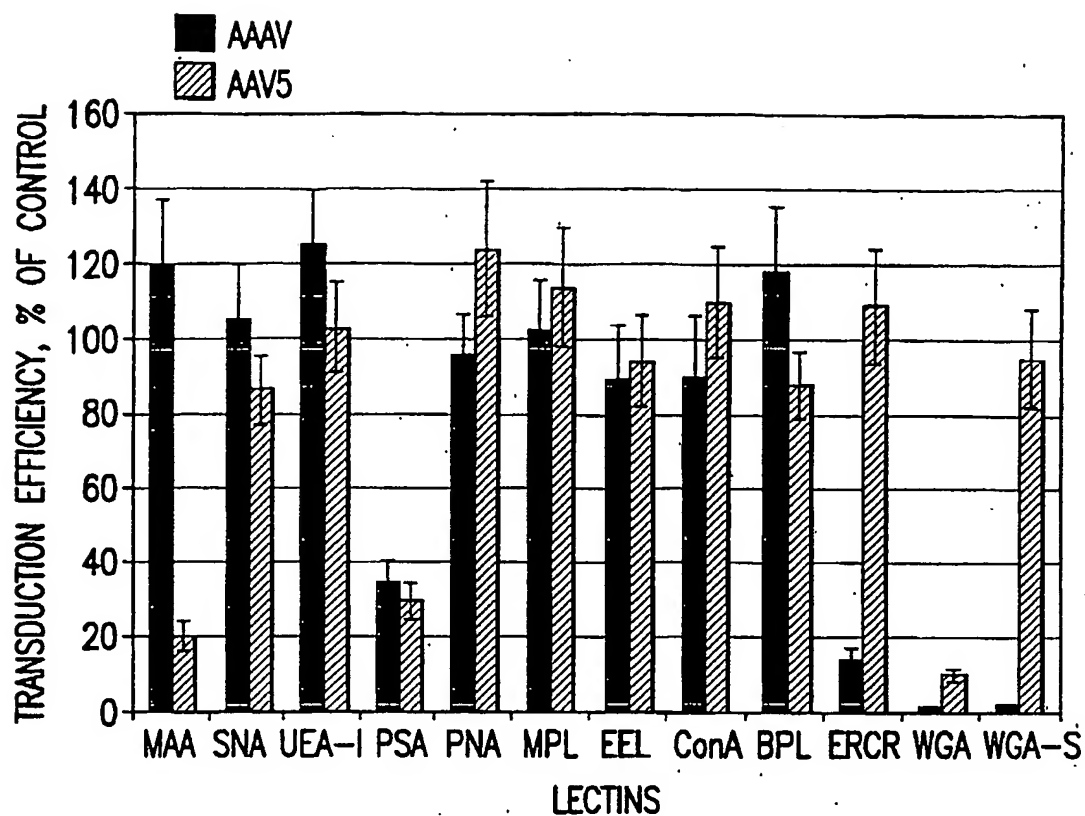


FIG.13A

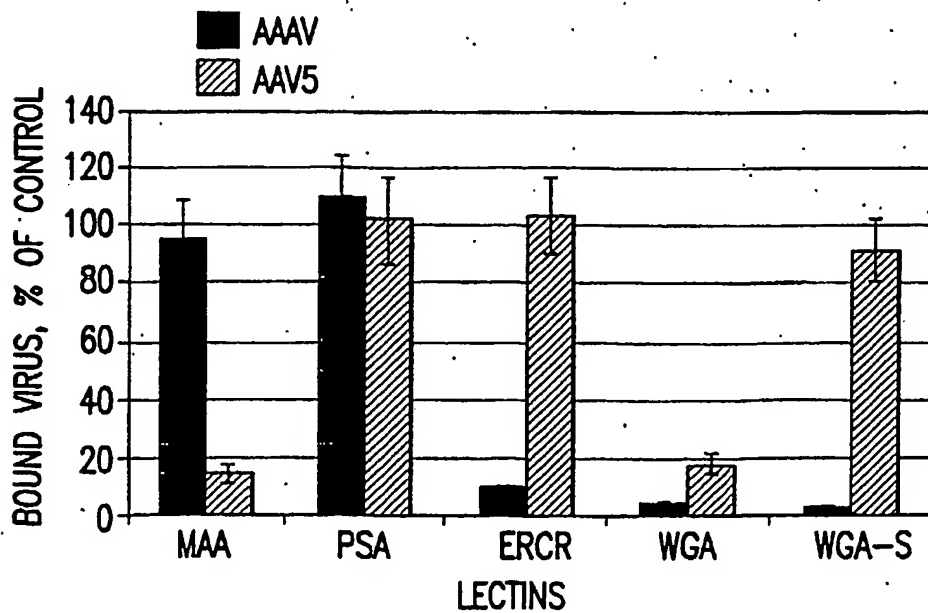


FIG.13B

## SEQUENCE LISTING

<110> Government of the United States of America, as represented by the  
Secretary, Department of Health & Human Services, c/o National  
Institutes  
of Health

<120> AVIAN ADENOASSOCIATED VIRUS (AAAV) AND  
USES THEREOF

<130> 14014.0412P1

<140> Unassigned

<141> 2004-05-18

<150> 60/472,066

<151> 2003-05-19

<160> 24

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4694

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 1

tggccagttt	ccaagacagg	ctcgcctcgt	cactcggggc	ggggcccaa	agggggccct	60
agcgaccgct	tcgcggtcgc	ggcccgagtg	agcgagcgag	cctgtcttgg	aaactggcca	120
gcactccggt	gaggtaatgc	cgtcacgtgg	tcgggaatgg	gaacgggaaa	tctcgcgaga	180
acgtaaacia	atataagacg	gcgccacacg	gcgctgcgtc	atacgcgcg	gcgcaccggc	240
gagatgaggt	cgtactacga	ggatcatcgt	cagctgcccc	acgacgtcga	gagtcaggta	300
cctggaatct	ccgattcgtt	cgtcaactgg	attacgtcgc	gagaatggac	gttgccctgag	360
gacgccgatt	gggatttggg	ccaggctcat	caagttcaac	tgacgctcgg	cgacaaaatc	420
caacggggaga	ttcgaactca	ttgggggacg	atggccaaag	aaccggactt	tcactatatt	480
atccaaactgg	aacaagggtg	ggtgttcttt	catttacacg	tcctgctgga	aacgtgttcc	540
gttaagccga	tgtgtactcg	aagatatatc	cgacatatct	aacaaaaaat	tgtgagtaaa	600
gtctactgcg	ccacgagcct	acgatggaag	gatggatgcg	tggtgaccaa	gacaaaaaat	660
ttcggggggcg	cgaacaagggt	ccggggccgag	tcgtatatct	ccgcctacct	gatcccgaaa	720
cagcaaccgg	aagtgcagtg	ggcgtggact	aacgtgcccc	agtatataaa	agcgtgcttg	780
caccgagaac	tgcgtgccag	tctcgcgcga	cttcacttcg	aggaggcggg	cgtctcgaac	840
tccaaggaaa	atctcgcgag	aactgcagac	ggcgctcccc	tgatgccgac	ccgcgtcagc	900
aaacgctaca	tggagctcgt	ggattggctc	gtggagaagg	ggatcaccac	cgagaaggaa	960
tggctgctgg	aaaacagaga	aagctttcgg	agctttcagg	cctcgagcaa	ctcggcgctg	1020
cagatcaaga	cggccctgca	aggcgccatt	caggagatgc	ttctgaccaa	gacggcggag	1080
gactacctcg	tcggaaagga	tccggtctcg	gacgacgaca	tcggtcagaa	ccgcatctac	1140
aagattctgg	aactgaacca	ctacgaccca	gcgtacgtgg	ggagtatttt	ggtcgggtgg	1200
tgccagaaga	aatgggggcaa	gcgaaacacg	ctgtggctgt	tcggacatgc	gaccaccggc	1260
aagaccaaca	tcgcgagggc	tattgcccct	gctgtgcccgt	tctatggatg	cgtaaactgg	1320
accaacgaga	actttccggt	caacgactgc	gtcgaaaaaa	tgattatctg	gtgggaggag	1380
ggcaaaatga	ccgccaaaagt	ggtggaaaca	gccaaaggcg	ttctggggagg	atctcgggtg	1440
agagtggacc	aaaaatgcaa	agcttcgggt	ccgatcgaa	cgacgccggg	cattattacc	1500
agtaacacca	acatgtgtta	tgcatcgac	gggaacacga	ccacgttcga	gcataagcag	1560
ccgttggagg	acaggatggt	taagctcgaa	ttgctgactc	ggttgccctga	tgactttggt	1620
aaggtgacca	aacaggaggt	gcgtcaattc	ttcaggtggt	ctcaggatca	cctgaccctt	1680

gtgatccag	aattcctagt	gcggaaggcg	gagtcctcgca	aaagaccgc	cccttcggg	1740
gaaggctata	taagcccgac	aaagcgccc	gcgctcgag	agcagcagca	ggcgtcggag	1800
agcgcgagcc	cggttcccac	caggtatcgt	atcaaatgct	cgaaacattg	cggatggat	1860
aaaatgtgt	ttccttgcca	aatttgtgaa	tcgatgaaca	gagatattaa	tatttgtgct	1920
attcataaaa	cgaccgactg	taaagagtgt	ttccccgact	acgggggataa	agatgatgta	1980
gaactacccc	cctgtacaga	acacaacgtg	tctcgttgtt	atcaatgtca	ttcggggcgaa	2040
ttgtatcgcg	tgacttcgga	ctctgacgag	aaacctgccc	ccgagagtga	tgaaggcacc	2100
gagccatcct	atgctccctg	cacgattcac	cacctgatgg	gcaagagtca	cgggttagtc	2160
acttgcgcg	cgtgtcggtt	gaaaaatagt	acgttgcag	atgacttggg	tgacggtgat	2220
ctcgaacaat	aaatgattga	aatgtagcca	tgtctctcat	ttctgatgcg	attccagatt	2280
ggttggagcg	gttgggtcaaa	aagggagtga	atgctgcagc	tgattttctac	catttggaaa	2340
gcggtcctcc	tcgtcctaag	gcaaatcagc	aaactcaaga	atctcttgaa	aaggacgatt	2400
cgagaggctc	cgtgttccca	ggctacaatt	atctaggccc	tttcaacggt	ctagataaag	2460
gagaacccgt	caacgaggca	gacgtgccc	ccttagaaca	cgacaaggct	tacgacctcg	2520
aaatcaaggga	cgggcacaaac	ccgtactttg	agtacaacga	ggccgacaga	cgtttccagg	2580
aacgtctcaa	agacgatacc	tcctttggag	gcaatttagg	taaagccatc	ttccaggcca	2640
aaaagagggg	tctcgaaccc	tttgggtctg	tggaagactc	aaagacggct	ccgaccggag	2700
acaagcgga	aggcgaagac	gaacctcggt	tgcccgacac	ttcttcacag	actccaaga	2760
aaaacaagaa	gcctcgcaag	gaaagacctt	ccggcggggc	agaagatccg	ggcgaaggca	2820
cctcttccaa	cgctggagca	gcagcaccgc	cctctagtgt	gggatcatct	atcatggctg	2880
aaggagggtg	cggcccagtg	ggcgatgcag	gccagggtgc	cgatggagtg	ggcaattcct	2940
ccggaaattg	gcattgcgat	tcccaatggc	tggaaaacgg	agtcgtcact	cgaaccaccc	3000
gaacctgggt	cttgcccagc	tacaacaacc	acctgtacaa	acgaatccaa	ggaccacgag	3060
gaggcgacaa	caacaacaaa	ttctttggat	tcagcaccgc	ctggggatac	tttgactaca	3120
atcgattcca	ctggcacttt	tcccgcgag	actggcaacg	actcatcaac	aacaactggg	3180
gcactcgtcc	caaagcgatg	cgcttttagac	tctttaacat	ccagggtaaa	gagggtcacgg	3240
tccaagactt	caacaccacc	atcggaacaa	acctcaccag	tacgggtccag	gtctttcgcg	3300
acaaggacta	ccaactgccg	tacgtcctcg	gatcggtac	cgaaggcacc	ttcccgcggt	3360
tcccagcgga	tatctacacg	atcccgcagt	acgggtactg	cacgctaaac	tacaacaacg	3420
aggcggtgga	tcgttcggcc	ttctactgtc	tggactactt	tccttcagac	atgctgcgga	3480
caggaaataa	ctttgagttt	acttacacct	tcgaggacgt	tcctttccat	agcatgtttg	3540
cccacaacca	gacgctagac	cggtgatgag	atcccctcgt	ggatcagtac	ctctgggctt	3600
tcagctccgt	cagccaagca	ggctcatctg	gacgaaactc	tcattactcg	cgggcgacta	3660
aaaccaacat	ggcggtctca	tataggaact	ggttacctgg	gcctttcttc	cgtgatcagc	3720
aaatctttac	ggcgctagc	aacatcacta	aaaataacgt	ctttagcggt	tgggaaaaag	3780
gcaagcaatg	ggaactcgac	aatcggacca	acctaatgca	gcccggctct	gcggcagcga	3840
ccacctttag	cggagaacct	gacctgcaag	ccatgcaaaa	cacgctgggt	tttagcagga	3900
ccgtctacga	tcaaacgacc	gccacgaccg	atcgtaacca	gatactcatc	accaacgaag	3960
acgaaatcag	accacccaac	tcggtcggta	tcgacgcgtg	gggagcagtt	cccaccaaca	4020
accagtcgat	cgtgaccccc	ggcactcgcg	cggccgtcaa	caatcaaggg	gcgcttcccg	4080
ggatgggtgtg	gcaaaacaga	gacatttacc	ctacagggac	ccatttggcc	aaaattcccg	4140
acactgacaa	tcacttccat	ccgtccccgc	ttattgggcg	gtttggctgc	aagcatcccc	4200
ctccccagat	tttcattaaa	aacacaccgc	tccttgccaa	cccttcggaa	acgttccaga	4260
cggccaaagt	ggcctccttc	atcaaccagt	actcgaccgg	acagtgcacc	gtcgaaatct	4320
tttgggaact	caagaaggaa	acctccaagc	gctggaaccc	cgaaatccag	ttcacctcca	4380
actttggcaa	cgggcccagc	atccagtttg	ccgtctccga	cacgggatcc	tattccgaac	4440
ctcgtcccat	cggtagccgt	taccttacca	aacctctgta	aattaaaccc	ttcaataaac	4500
cgtttatgcg	taactgtatt	tcggtctcct	gtcggttattc	agtcacatga	tgcggcatta	4560
cctcaccgga	gtgctggcca	gtttccaaga	caggctcgtc	cgctcactcg	ggccggggcc	4620
ccaaaggggc	ccctagcgac	cgcttcgcgg	tcgcgggccc	agtgagcgag	cgagcctgtc	4680
ttggaaactg	gccaa					4694

&lt;210&gt; 2

&lt;211&gt; 1989

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 2

```

atgaggtcgt actacgaggt catcgttcag ctgcccacg acgtcgagag tcaggtacct      60
ggaatctccg attcgttcgt caactggatt acgtcgcgag aatggacgtt gcctgaggac      120
gccgattggg atttggacca ggtcgatcaa gttcaactga cgctcggcga caaaatccaa      180
cgggagattc gaactcattg ggggacgatg gccaaagaac cggactttca ctattttatc      240
caactcggaac aaggtgaggt gttctttcat ttacacgtcc tgctggaaac gtgttccgta      300
aagccgatgg tactcggaag atatatccga catattcaac aaaaaattgt gagtaaagtc      360
tactgcgcca cgagcctacg atggaaggat ggatgcgtgg tgaccaagac caaaaatttc      420
gggggcgcga acaagggtccg ggccgagtcg tatattcccc cctacctgat cccgaaacag      480
caaccggaag tgcagtgggc gtggactaac gtgcccagat atataaaagc gtgcttgcac      540
cgagaactgc gtgccagtct cgcgcgactt cacttcgagg aggcgggcgt ctcgcaatcc      600
aaggaataac tcgcgagaac tgcagacggc gctccgtga tgccgaccg cgtcagcaaa      660
cgctacatgg agctcgtgga ttggctcgtg gagaagggga tcaccaccga gaaggaaatgg      720
ctgctggaaa acagagaaaag ctttcggagc ttccaggcct cgagcaactc ggcgcgtcag      780
atcaagacgg cctgcaagg cgccattcag gagatgcttc tgaccaagac ggcggaggac      840
tacctcgtcg gaaaggatcc cgtctcggac gacgacatcc gtcagaaccg catctacaag      900
attctggaac tgaaccacta cgaccacagc tacgtgggga gtattttggt cgggtggtgc      960
cagaagaaat ggggcaagcg aaacacgctg tggctgttcg gacatgcgac caccggaag      1020
accaacatcg cggaggctat tgccatgct gtgcccgttc atggatgcgt taactggacc      1080
aacgagaact ttccgttcaa cgactcgtc gaaaaaatga ttatctggtg ggaggagggc      1140
aaaatgaccg ccaaagtggg ggaacagcc aaggcgattc tgggaggatc tcgggtgaga      1200
gtggacaaa aatgcaaagc ttcggttcg atcgaaacga cgccggtcat tattaccagt      1260
aacaccaaca tgtgttatgt catcgacggg aacacgacca cgttcgagca taagcagccg      1320
ttggaggaca ggatgtttaa gctcgaattg ctgactcggg tgccctgatga ctttggtaa      1380
gtgacaaac aggaggtgcg tcaattcttc aggtggtctc aggatcacct gaccctgtg      1440
atcccagaat tcctagtgcg gaagcgagg tctcgcaaaa gaccgcgcc ttccggggaa      1500
ggctatataa gcccgacaaa gcggcccgcg ctcgacagag agcagcaggc gtcggagagc      1560
gcggaccggg ttccaccag gtatcgtatc aaatgctcga aacattgcgg tatggataaa      1620
atgttggttc cttgccaaat ttgtgaatcg atgaacagag atattaatat ttgtgctatt      1680
cataaaacga ccgactgtaa agagtgttc cccgactacg gggataaaga tgatgtagaa      1740
ctacccccct gtacagaaca caacgtgtct cgttgttatc aatgtcattc gggcgaattg      1800
tatecgtgta cttcggactc tgacgagaaa cctgcccccg agagtgatga aggcaccgag      1860
ccatcctatg ctccctgcac gattcaccac ctgatgggca agagtcacgg gttagtcact      1920
tgcgcgcggt gtcggttgaa aaatagtacg ttgcatgatg acttgatga cggatgatc      1980
gaacaataa

```

&lt;210&gt; 3

&lt;211&gt; 662

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 3

```

Met Arg Ser Tyr Tyr Glu Val Ile Val Gln Leu Pro Asn Asp Val Glu
1           5           10           15
Ser Gln Val Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Ile Thr Ser
20           25           30
Arg Glu Trp Thr Leu Pro Glu Asp Ala Asp Trp Asp Leu Asp Gln Val
35           40           45
Asp Gln Val Gln Leu Thr Leu Gly Asp Lys Ile Gln Arg Glu Ile Arg
50           55           60
Thr His Trp Gly Thr Met Ala Lys Glu Pro Asp Phe His Tyr Phe Ile
65           70           75           80
Gln Leu Glu Gln Gly Glu Val Phe Phe His Leu His Val Leu Leu Glu
85           90           95
Thr Cys Ser Val Lys Pro Met Val Leu Gly Arg Tyr Ile Arg His Ile
100           105           110

```

Gln Gln Lys Ile Val Ser Lys Val Tyr Cys Ala Thr Ser Leu Arg Trp  
           115                          120                          125  
 Lys Asp Gly Cys Val Val Thr Lys Thr Lys Asn Phe Gly Gly Ala Asn  
           130                          135                          140  
 Lys Val Arg Ala Glu Ser Tyr Ile Pro Ala Tyr Leu Ile Pro Lys Gln  
           145                          150                          155                          160  
 Gln Pro Glu Val Gln Trp Ala Trp Thr Asn Val Pro Glu Tyr Ile Lys  
                           165                          170                          175  
 Ala Cys Leu His Arg Glu Leu Arg Ala Ser Leu Ala Arg Leu His Phe  
                           180                          185                          190  
 Glu Glu Ala Gly Val Ser Gln Ser Lys Glu Asn Leu Ala Arg Thr Ala  
           195                          200                          205  
 Asp Gly Ala Pro Val Met Pro Thr Arg Val Ser Lys Arg Tyr Met Glu  
           210                          215                          220  
 Leu Val Asp Trp Leu Val Glu Lys Gly Ile Thr Thr Glu Lys Glu Trp  
           225                          230                          235                          240  
 Leu Leu Glu Asn Arg Glu Ser Phe Arg Ser Phe Gln Ala Ser Ser Asn  
                           245                          250                          255  
 Ser Ala Arg Gln Ile Lys Thr Ala Leu Gln Gly Ala Ile Gln Glu Met  
           260                          265                          270  
 Leu Leu Thr Lys Thr Ala Glu Asp Tyr Leu Val Gly Lys Asp Pro Val  
           275                          280                          285  
 Ser Asp Asp Asp Ile Arg Gln Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
           290                          295                          300  
 Asn His Tyr Asp Pro Ala Tyr Val Gly Ser Ile Leu Val Gly Trp Cys  
           305                          310                          315                          320  
 Gln Lys Lys Trp Gly Lys Arg Asn Thr Leu Trp Leu Phe Gly His Ala  
                           325                          330                          335  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro  
           340                          345                          350  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
           355                          360                          365  
 Cys Val Glu Lys Met Ile Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
           370                          375                          380  
 Lys Val Val Glu Thr Ala Lys Ala Ile Leu Gly Gly Ser Arg Val Arg  
           385                          390                          395                          400  
 Val Asp Gln Lys Cys Lys Ala Ser Val Pro Ile Glu Pro Thr Pro Val  
                           405                          410                          415  
 Ile Ile Thr Ser Asn Thr Asn Met Cys Tyr Val Ile Asp Gly Asn Thr  
           420                          425                          430  
 Thr Thr Phe Glu His Lys Gln Pro Leu Glu Asp Arg Met Phe Lys Leu  
           435                          440                          445  
 Glu Leu Leu Thr Arg Leu Pro Asp Asp Phe Gly Lys Val Thr Lys Gln  
           450                          455                          460  
 Glu Val Arg Gln Phe Phe Arg Trp Ser Gln Asp His Leu Thr Pro Val  
           465                          470                          475                          480  
 Ile Pro Glu Phe Leu Val Arg Lys Ala Glu Ser Arg Lys Arg Pro Ala  
                           485                          490                          495  
 Pro Ser Gly Glu Gly Tyr Ile Ser Pro Thr Lys Arg Pro Ala Leu Ala  
           500                          505                          510  
 Glu Gln Gln Gln Ala Ser Glu Ser Ala Asp Pro Val Pro Thr Arg Tyr  
           515                          520                          525  
 Arg Ile Lys Cys Ser Lys His Cys Gly Met Asp Lys Met Leu Phe Pro  
           530                          535                          540  
 Cys Gln Ile Cys Glu Ser Met Asn Arg Asp Ile Asn Ile Cys Ala Ile  
           545                          550                          555                          560  
 His Lys Thr Thr Asp Cys Lys Glu Cys Phe Pro Asp Tyr Gly Asp Lys  
                           565                          570                          575  
 Asp Asp Val Glu Leu Pro Pro Cys Thr Glu His Asn Val Ser Arg Cys  
           580                          585                          590



Tyr Gln Cys His Ser Gly Glu Leu Tyr Arg Val Thr Ser Asp Ser Asp  
           595                                600                                605  
 Glu Lys Pro Ala Pro Glu Ser Asp Glu Gly Thr Glu Pro Ser Tyr Ala  
           610                                615                                620  
 Pro Cys Thr Ile His His Leu Met Gly Lys Ser His Gly Leu Val Thr  
           625                                630                                635                                640  
 Cys Ala Ala Cys Arg Leu Lys Asn Ser Thr Leu His Asp Asp Leu Asp  
                                 645                                650                                655  
 Asp Gly Asp Leu Glu Gln  
                                 660

<210> 4  
 <211> 1323  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence; note =  
           synthetic construct

<400> 4  
 atggagctcg tggattggct cgtggagaag gggatcacca ccgagaagga atggctgctg 60  
 gaaaacagag aaagctttcg gagctttcag gcctcgagca actcggcgcg tcagatcaag 120  
 acggccctgc aaggcgccat tcaggagatg cttctgacca agacggcgga ggactacctc 180  
 gtcggaagg atcccgctctc ggacgacgac atccgtcaga accgcatcta caagattctg 240  
 gaactgaacc actacgaccc agcgtacgtg gggagtattt tggctcgggtg gtgccagaag 300  
 aaatggggca agcgaaacac gctgtggctg ttcggacatg cgaccaccgg caagaccaac 360  
 atcgcggagg ctattgcccc tgcgtgccc ttcctatggat gcgttaactg gaccaacgag 420  
 aactttccgt tcaacgactg cgtcgaaaaa atgattatct ggtgggagga gggcaaaatg 480  
 accgccaag tggtggaac agccaaggcg attctgggag gatctcgggt gagagtggac 540  
 caaaaatgca aagcttcggt tccgatcgaa ccgacgccgg tcattattac cagtaacacc 600  
 aacatgtgtt atgtcatcga cgggaacacg accacgttcg agcataagca gccgttgag 660  
 gacaggatgt ttaagctcga attcgtgact cggttgcctg atgactttgg taaggtgacc 720  
 aaacaggagg tgcgtcaatt cttcagggtg tctcaggatc acctgacccc tgtgatccca 780  
 gaattcctag tgcggaaggc ggagtctcgc aaaagaccgg ccccttcggg ggaaggctat 840  
 ataagcccga caaagcggcc cgcgctcgca gagcagcagc aggcgtcgga gagcgcggac 900  
 ccggttcccc ccagggtatcg tatcaaatgc tcgaaacatt gcggtatgga taaaatgttg 960  
 tttccttgcc aaatttgtga atcgatgaac agagatatta atatttgtgc tattcataaa 1020  
 acgaccgact gtaaagagtg tttccccgac tacggggata aagatgatgt agaactaccc 1080  
 cctgtgtacg aacacaacgt gtctcgttgt tatcaatgtc attcgggcga attgtatcgc 1140  
 gtgaacttcgg actctgacga gaaacctgcc cccgagagtg atgaaggcac cgagccatcc 1200  
 tatgtccct gcacgattca ccacctgatg ggcaagagtc acgggttagt cacttgcgcg 1260  
 gcgtgtcggg tgaaaaatag tacgttgcac gatgacttgg atgacggtga tctcgaacaa 1320  
 taa 1323

<210> 5  
 <211> 440  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence; note =  
           synthetic construct

<400> 5  
 Met Glu Leu Val Asp Trp Leu Val Glu Lys Gly Ile Thr Thr Glu Lys  
   1                                5                                10                                15  
 Glu Trp Leu Leu Glu Asn Arg Glu Ser Phe Arg Ser Phe Gln Ala Ser  
           20                                25                                30  
 Ser Asn Ser Ala Arg Gln Ile Lys Thr Ala Leu Gln Gly Ala Ile Gln  
           35                                40                                45

Glu Met Leu Leu Thr Lys Thr Ala Glu Asp Tyr Leu Val Gly Lys Asp  
 50 55 60  
 Pro Val Ser Asp Asp Ile Arg Gln Asn Arg Ile Tyr Lys Ile Leu  
 65 70 75 80  
 Glu Leu Asn His Tyr Asp Pro Ala Tyr Val Gly Ser Ile Leu Val Gly  
 85 90 95  
 Trp Cys Gln Lys Lys Trp Gly Lys Arg Asn Thr Leu Trp Leu Phe Gly  
 100 105 110  
 His Ala Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala  
 115 120 125  
 Val Pro Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe  
 130 135 140  
 Asn Asp Cys Val Glu Lys Met Ile Ile Trp Trp Glu Glu Gly Lys Met  
 145 150 155 160  
 Thr Ala Lys Val Val Glu Thr Ala Lys Ala Ile Leu Gly Gly Ser Arg  
 165 170 175  
 Val Arg Val Asp Gln Lys Cys Lys Ala Ser Val Pro Ile Glu Pro Thr  
 180 185 190  
 Pro Val Ile Ile Thr Ser Asn Thr Asn Met Cys Tyr Val Ile Asp Gly  
 195 200 205  
 Asn Thr Thr Thr Phe Glu His Lys Gln Pro Leu Glu Asp Arg Met Phe  
 210 215 220  
 Lys Leu Glu Leu Leu Thr Arg Leu Pro Asp Asp Phe Gly Lys Val Thr  
 225 230 235 240  
 Lys Gln Glu Val Arg Gln Phe Phe Arg Trp Ser Gln Asp His Leu Thr  
 245 250 255  
 Pro Val Ile Pro Glu Phe Leu Val Arg Lys Ala Glu Ser Arg Lys Arg  
 260 265 270  
 Pro Ala Pro Ser Gly Glu Gly Tyr Ile Ser Pro Thr Lys Arg Pro Ala  
 275 280 285  
  
 Leu Ala Glu Gln Gln Gln Ala Ser Glu Ser Ala Asp Pro Val Pro Thr  
 290 295 300  
 Arg Tyr Arg Ile Lys Cys Ser Lys His Cys Gly Met Asp Lys Met Leu  
 305 310 315 320  
 Phe Pro Cys Gln Ile Cys Glu Ser Met Asn Arg Asp Ile Asn Ile Cys  
 325 330 335  
 Ala Ile His Lys Thr Thr Asp Cys Lys Glu Cys Phe Pro Asp Tyr Gly  
 340 345 350  
 Asp Lys Asp Asp Val Glu Leu Pro Pro Cys Thr Glu His Asn Val Ser  
 355 360 365  
 Arg Cys Tyr Gln Cys His Ser Gly Glu Leu Tyr Arg Val Thr Ser Asp  
 370 375 380  
 Ser Asp Glu Lys Pro Ala Pro Glu Ser Asp Glu Gly Thr Glu Pro Ser  
 385 390 395 400  
 Tyr Ala Pro Cys Thr Ile His His Leu Met Gly Lys Ser His Gly Leu  
 405 410 415  
 Val Thr Cys Ala Ala Cys Arg Leu Lys Asn Ser Thr Leu His Asp Asp  
 420 425 430  
 Leu Asp Asp Gly Asp Leu Glu Gln  
 435 440

&lt;210&gt; 6

&lt;211&gt; 1626

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

 <223> Description of Artificial Sequence; note =  
 synthetic construct

&lt;400&gt; 6

```

atgaggtcgt actacgaggt catcgttcag ctgcccacg acgtcgagag tcaggtacct      60
ggaatctccg attcgttcgt caactggatt acgtcgcgag aatggacgtt gcctgaggac      120
gccgattggg atttggaacca ggtcgatcaa gttcaactga cgctcggcga caaaatccaa      180
cgggagattc gaactcattg ggggacgatg gccaaagaac cggactttca ctattttatc      240
caactggaac aaggtgaggt gttctttcat ttacacgtcc tgctggaaac gtgttccgta      300
aagccgatgg tactcggaag atatatccga catattcaac aaaaaattgt gagtaaagtc      360
tactgcgcca cgagcctacg atggaaggat ggatgcgtgg tgaccaagac caaaaatttc      420
gggggcgcca acaaggtccg ggccgagtcg tatattccc cctacctgat cccgaaacag      480
caaccggaag tgcagtgggc gtggactaac gtgcccagat atataaaagc gtgcttgacac      540
cgagaactgc gtgccagtct cgcgcgactt cacttcgagg aggcgggctg ctcgcaatcc      600
aaggaaaatc tcgcgagaac tgcagacggc gctcccgtga tgccgaccgc cgctcagcaa      660
cgctacatgg agctcgtgga ttggctcgtg gagaagggga tcaccaccga gaaggaatgg      720
ctgctggaaa ccaagaaaag ctttcggagc tttcaggcct cgagcaactc ggccgctcag      780
atcaagacgg ccctgcaaag cgccattcag gagatgcttc tgaccaagac ggccgaggac      840
tacctcgtcg gaaaggatcc cgtctcggac gacgacatcc gtcagaaccg catctacaag      900
attctggaac tgaaccacta cgaccacgag tacgtgggga gtatttttgt cgggtggtgc      960
cagaagaaat ggggcaagcg aaacacgctg tggtctgtcg gacatgcgac caccggcaag     1020
accaacatcg cggaggctat tgcccagctt gtgcccgttc atggatgcgt taactggacc     1080
aacgagaact ttccgttcaa cgactgcgtc gaaaaaatga ttatctggtg ggaggagggc     1140
aaaatgaccg ccaaagtgtt ggaacagccc aaggcgattc tgggaggatc tcgggtgaga     1200
gtggaccaaa aatgcaaacg ttccggttccg atcgaaaccg cgccggtcat tattaccagt     1260
aacaccaaca tgtgttatgt catcgacggg aacacgacca cgttcgagca taagcagccg     1320
ttggaggaca ggatgtttaa gctcgaattg ctgactcggg tgccgatga ctttggttaag     1380
gtgaccaaac aggaggtgcg tcaattcttc aggtggtctc aggatcacct gaccctgtg      1440
atcccagaat tcctagtgcg gaaggcggag tctcgcaaaa gaccgcgcc ttccggggaa      1500
ggctatataa gcccgcacaa gcggcccgcg ctgcgagagc agcagcaggc gtcggagagc      1560
gcggaccggg ttcccaccag attggttggg gcggttggtc aaaaaggag tgaatgctgc      1620
agctga

```

&lt;210&gt; 7

&lt;211&gt; 541

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 7

```

Met Arg Ser Tyr Tyr Glu Val Ile Val Gln Leu Pro Asn Asp Val Glu
 1              5              10              15
Ser Gln Val Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Ile Thr Ser
      20              25              30
Arg Glu Trp Thr Leu Pro Glu Asp Ala Asp Trp Asp Leu Asp Gln Val
      35              40              45
Asp Gln Val Gln Leu Thr Leu Gly Asp Lys Ile Gln Arg Glu Ile Arg
      50              55              60
Thr His Trp Gly Thr Met Ala Lys Glu Pro Asp Phe His Tyr Phe Ile
      65              70              75              80
Gln Leu Glu Gln Gly Glu Val Phe Phe His Leu His Val Leu Leu Glu
      85              90              95
Thr Cys Ser Val Lys Pro Met Val Leu Gly Arg Tyr Ile Arg His Ile
      100             105             110
Gln Gln Lys Ile Val Ser Lys Val Tyr Cys Ala Thr Ser Leu Arg Trp
      115             120             125
Lys Asp Gly Cys Val Val Thr Lys Thr Lys Asn Phe Gly Gly Ala Asn
      130             135             140
Lys Val Arg Ala Glu Ser Tyr Ile Pro Ala Tyr Leu Ile Pro Lys Gln
      145             150             155             160

```

Gln Pro Glu Val Gln Trp Ala Trp Thr Asn Val Pro Glu Tyr Ile Lys  
 Ala Cys Leu His Arg Glu Leu Arg Ala Ser Leu Ala Arg Leu His Phe  
 Glu Glu Ala Gly Val Ser Gln Ser Lys Glu Asn Leu Ala Arg Thr Ala  
 Asp Gly Ala Pro Val Met Pro Thr Arg Val Ser Lys Arg Tyr Met Glu  
 Leu Val Asp Trp Leu Val Glu Lys Gly Ile Thr Thr Glu Lys Glu Trp  
 Leu Leu Glu Asn Arg Glu Ser Phe Arg Ser Phe Gln Ala Ser Ser Asn  
 Ser Ala Arg Gln Ile Lys Thr Ala Leu Gln Gly Ala Ile Gln Glu Met  
 Leu Leu Thr Lys Thr Ala Glu Asp Tyr Leu Val Gly Lys Asp Pro Val  
 Ser Asp Asp Asp Ile Arg Gln Asn Arg Ile Tyr Lys Ile Leu Glu Leu  
 Asn His Tyr Asp Pro Ala Tyr Val Gly Ser Ile Leu Val Gly Trp Cys  
 Gln Lys Lys Trp Gly Lys Arg Asn Thr Leu Trp Leu Phe Gly His Ala  
 Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro  
 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp  
 Cys Val Glu Lys Met Ile Ile Trp Trp Glu Glu Gly Lys Met Thr Ala  
 Lys Val Val Glu Thr Ala Lys Ala Ile Leu Gly Gly Ser Arg Val Arg  
 Val Asp Gln Lys Cys Lys Ala Ser Val Pro Ile Glu Pro Thr Pro Val  
 Ile Ile Thr Ser Asn Thr Asn Met Cys Tyr Val Ile Asp Gly Asn Thr  
 Thr Thr Phe Glu His Lys Gln Pro Leu Glu Asp Arg Met Phe Lys Leu  
 Glu Leu Leu Thr Arg Leu Pro Asp Asp Phe Gly Lys Val Thr Lys Gln  
 Glu Val Arg Gln Phe Phe Arg Trp Ser Gln Asp His Leu Thr Pro Val  
 Ile Pro Glu Phe Leu Val Arg Lys Ala Glu Ser Arg Lys Arg Pro Ala  
 Pro Ser Gly Glu Gly Tyr Ile Ser Pro Thr Lys Arg Pro Ala Leu Ala  
 Glu Gln Gln Gln Ala Ser Glu Ser Ala Asp Pro Val Pro Thr Arg Leu  
 Val Gly Ala Val Gly Gln Lys Gly Ser Glu Cys Cys Ser  
 165 170 175  
 180 185 190  
 195 200 205  
 210 215 220  
 225 230 235  
 245 250 255  
 260 265 270  
 275 280 285  
 290 295 300  
 305 310 315  
 325 330 335  
 340 345 350  
 355 360 365  
 370 375 380  
 385 390 395  
 405 410 415  
 420 425 430  
 435 440 445  
 450 455 460  
 465 470 475  
 485 490 495  
 500 505 510  
 515 520 525  
 530 535 540

&lt;210&gt; 8

&lt;211&gt; 960

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

 <223> Description of Artificial Sequence; note =  
 synthetic construct

&lt;400&gt; 8

```

atggagctcg tggattggct cgtggagaag gggatcacca ccgagaagga atggctgctg      60
gaaaacagag aaagctttcg gagctttcag gcctcgagca actcggcgcg tcagatcaag      120
acggccctgc aaggcgccat tcaggagatg cttctgacca agacggcgga ggactacctc      180
gtcggaaaagg atcccgctc ggacgacgac atccgtcaga accgcatcta caagattctg      240
gaactgaacc actacgaccc agcgtacgtg gggagtatTT tggtcgggtg gtgccagaag      300
aaatggggca agcgaacac gctgtggctg ttcggacatg cgaccaccgg caagaccaac      360
atcgcgagg ctattgccca tgctgtgccg ttctatggat gcgttaactg gaccaacgag      420
aactttccgt tcaacgactg cgtcgaaaaa atgattatct ggtgggagga gggcaaaatg      480
accgccaaag tgggtgaaac agccaaggcg attctgggag gatctcgggt gagagtggac      540
caaaaatgca aagcttcggt tccgatcgaa ccgacgccgg tcattattac cagtaacacc      600
aacatgtggt atgtcatcga cgggaacacg accacgttcg agcataagca gccgttgag      660
gacaggatgt ttaagctcga attgtgact cggttcctg atgactttgg taaggtagc      720
aaacaggagg tgcgtcaatt cttcagggtg tctcaggatc acctgacccc tgtgatccca      780
gaattcctag tgcggaaggc ggagtctcgc aaaagacccg ccccttcggg ggaaggctat      840
ataagcccga caaagcgcc cgcgctcgca gagcagcagc aggcgtcgga gagcgcgag      900
ccggttccca ccagattggt tggagcgggt ggtcaaaaaa ggagtgaatg ctgcagctga      960

```

&lt;210&gt; 9

&lt;211&gt; 319

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 9

```

Met Glu Leu Val Asp Trp Leu Val Glu Lys Gly Ile Thr Thr Glu Lys
 1          5          10         15
Glu Trp Leu Leu Glu Asn Arg Glu Ser Phe Arg Ser Phe Gln Ala Ser
 20         25         30
Ser Asn Ser Ala Arg Gln Ile Lys Thr Ala Leu Gln Gly Ala Ile Gln
 35         40         45
Glu Met Leu Leu Thr Lys Thr Ala Glu Asp Tyr Leu Val Gly Lys Asp
 50         55         60
Pro Val Ser Asp Asp Asp Ile Arg Gln Asn Arg Ile Tyr Lys Ile Leu
 65         70         75         80
Glu Leu Asn His Tyr Asp Pro Ala Tyr Val Gly Ser Ile Leu Val Gly
 85         90         95
Trp Cys Gln Lys Lys Trp Gly Lys Arg Asn Thr Leu Trp Leu Phe Gly
100        105        110
His Ala Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala
115        120        125
Val Pro Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe
130        135        140
Asn Asp Cys Val Glu Lys Met Ile Ile Trp Trp Glu Glu Gly Lys Met
145        150        155        160
Thr Ala Lys Val Val Glu Thr Ala Lys Ala Ile Leu Gly Gly Ser Arg
165        170        175
Val Arg Val Asp Gln Lys Cys Lys Ala Ser Val Pro Ile Glu Pro Thr
180        185        190
Pro Val Ile Ile Thr Ser Asn Thr Asn Met Cys Tyr Val Ile Asp Gly
195        200        205
Asn Thr Thr Thr Phe Glu His Lys Gln Pro Leu Glu Asp Arg Met Phe
210        215        220
Lys Leu Glu Leu Leu Thr Arg Leu Pro Asp Asp Phe Gly Lys Val Thr
225        230        235        240
Lys Gln Glu Val Arg Gln Phe Phe Arg Trp Ser Gln Asp His Leu Thr
245        250        255

```

Pro Val Ile Pro Glu Phe Leu Val Arg Lys Ala Glu Ser Arg Lys Arg  
 260 265 270  
 Pro Ala Pro Ser Gly Glu Gly Tyr Ile Ser Pro Thr Lys Arg Pro Ala  
 275 280 285  
 Leu Ala Glu Gln Gln Gln Ala Ser Glu Ser Ala Asp Pro Val Pro Thr  
 290 295 300  
 Arg Leu Val Gly Ala Val Gly Gln Lys Gly Ser Glu Cys Cys Ser  
 305 310 315

&lt;210&gt; 10

&lt;211&gt; 2232

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 10

```

atgtctctca tttctgatgc gattccagat tggttggagc ggttgggtcaa aaagggagtg      60
aatgtctcag ctgattttcta ccatttggaa agcggtcctc ctctgtcctaa ggcaaatacag      120
caaactcaag aatctcttga aaaggacgat tcgagagggtc tcgtgttccc aggctacaat      180
tatctaggcc ctttcaacgg tctagataaa ggagaacccg tcaacgaggc agacgctgcc      240
gccttagaac acgacaaggc ttacgacctc gaaatcaagg acgggcacaa cccgtacttt      300
gagtacaacg aggccgacag acgtttccag gaacgtctca aagacgatac ctcttttggg      360
ggcaatttag gtaaaagccat cttccaggcc aaaaagaggg ttctcgaacc ctttgggtctg      420
gtggaagact caaagacggc tccgaccgga gacaagcggg aaggcgaaga cgaacctcgt      480
ttgcccagaca cttcttcaca gactcccaag aaaaacaaga agcctcgcaa ggaaagacct      540
tccggcgggg cagaagatcc gggcgaaggc acctcttcca acgctggagc agcagcacc      600
gcctctagtg tgggatcatc tatcatggct gaaggaggtg gcggcccagt gggcgatgca      660
ggccagggtg ccgatggagt gggcaattcc tccggaattt ggcattgcga ttcccaatgg      720
ctggaataac gagtcgtcac tcgaaccacc cgaacctggg tcttgcccag ctacaacaac      780
cacctgtaca aacgaatcca aggaccagc ggaggcgaca acaacaaca attctttgga      840
ttcagcacc cctgggggata ctttgactac aatcgattcc actgccactt tccccgcga      900
gactggcaac gactcatcaa caacaactgg ggcattccgt ccaaagcgat gcgctttaga      960
ctctttaaca tccagggtta agaggtcacg gtccaagact tcaacaccac catcggaac      1020
aacctcacca gtacgttcca ggtctttgcg gacaaggact accaactgcc gtacgtcctc      1080
ggatcggtca ccgaaggcac cttcccgcgg tccccagcgg atatctacac gatcccgcag      1140
tacgggtact gcacgtctaa ctacaacaac gaggcgttgg atcggttcggc cttctactgt      1200
ctggactact ttccctcaga catgtgcgg acaggaata actttgagtt tacttacacc      1260
ttcaggagcg ttcccttcca tagcatgttt gccacaacc agacgctaga ccggctgatg      1320
aatcccctcg tggatcagta cctctgggct ttcagctccg tcagccaagc aggtcatct      1380
ggacgagctc ttcattactc gcgggcgact aaaaccaaca tggcgggtca atataggaac      1440
tggttacctg ggcctttctt ccgtgatcag caaatcttta cgggcgctag caacatcact      1500
aaaaataacg tctttagcgt ttgggaaaaa ggcaagcaat gggaactcga caatcggacc      1560
aacctaatac agcccggtc tgccggcagc accaccttta gcggagaacc tgaccgtcaa      1620
gccatgcaaa acacgttggc ttttagcagg accgtctacg atcaaacgac cgccacgacc      1680
gatcgtaacc agatactcat caccaacgaa gacgaaatca gaccaccaa ctcggtcggg      1740
atcgacgcgt ggggagcagt tcccaccaac aaccagtcga tcgtgacccc cggcactcgc      1800
gcggcgtcga acaatcaagg ggcgcttccc gggatggtgt ggcaaacag agacatttac      1860
cctacaggga cccatttggc caaaattccc gacactgaca atcacttcca tccgtccccg      1920
cttattgggc ggtttggctg caagcatccc cctcccaga ttttcattaa aaacacacc      1980
gtccctgcca acccttcgga aacgttccag acggccaaag tggcctcctt catcaaccag      2040
tactcgaccg gacagtgcac cgtcgaaatc ttttgggaac tcaagaagga aacctccaag      2100
cgctggaacc ccgaaatcca gttcacctcc aactttggca acgcggccga catccagttt      2160
gccgtctccg acacgggata ctattccgaa cctcgtccca tcggtaccgc ttaccttacc      2220
aaacctctgt aa                                     2232

```

&lt;210&gt; 11

&lt;211&gt; 743

&lt;212&gt; PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 11

```

Met Ser Leu Ile Ser Asp Ala Ile Pro Asp Trp Leu Glu Arg Leu Val
 1           5           10           15
Lys Lys Gly Val Asn Ala Ala Ala Asp Phe Tyr His Leu Glu Ser Gly
      20           25           30
Pro Pro Arg Pro Lys Ala Asn Gln Thr Gln Glu Ser Leu Glu Lys
      35           40           45
Asp Asp Ser Arg Gly Leu Val Phe Pro Gly Tyr Asn Tyr Leu Gly Pro
      50           55           60
Phe Asn Gly Leu Asp Lys Gly Glu Pro Val Asn Glu Ala Asp Ala Ala
      65           70           75           80
Ala Leu Glu His Asp Lys Ala Tyr Asp Leu Glu Ile Lys Asp Gly His
      85           90           95
Asn Pro Tyr Phe Glu Tyr Asn Glu Ala Asp Arg Arg Phe Gln Glu Arg
      100          105          110
Leu Lys Asp Asp Thr Ser Phe Gly Gly Asn Leu Gly Lys Ala Ile Phe
      115          120          125
Gln Ala Lys Lys Arg Val Leu Glu Pro Phe Gly Leu Val Glu Asp Ser
      130          135          140
Lys Thr Ala Pro Thr Gly Asp Lys Arg Lys Gly Glu Asp Glu Pro Arg
      145          150          155          160
Leu Pro Asp Thr Ser Ser Gln Thr Pro Lys Lys Asn Lys Lys Pro Arg
      165          170          175
Lys Glu Arg Pro Ser Gly Gly Ala Glu Asp Pro Gly Glu Gly Thr Ser
      180          185          190
Ser Asn Ala Gly Ala Ala Ala Pro Ala Ser Ser Val Gly Ser Ser Ile
      195          200          205

Met Ala Glu Gly Gly Gly Gly Pro Val Gly Asp Ala Gly Gln Gly Ala
      210          215          220
Asp Gly Val Gly Asn Ser Ser Gly Asn Trp His Cys Asp Ser Gln Trp
      225          230          235          240
Leu Glu Asn Gly Val Val Thr Arg Thr Thr Arg Thr Trp Val Leu Pro
      245          250          255
Ser Tyr Asn Asn His Leu Tyr Lys Arg Ile Gln Gly Pro Ser Gly Gly
      260          265          270
Asp Asn Asn Asn Lys Phe Phe Gly Phe Ser Thr Pro Trp Gly Tyr Phe
      275          280          285
Asp Tyr Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg
      290          295          300
Leu Ile Asn Asn Asn Trp Gly Ile Arg Pro Lys Ala Met Arg Phe Arg
      305          310          315          320
Leu Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Phe Asn Thr
      325          330          335
Thr Ile Gly Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ala Asp Lys
      340          345          350
Asp Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala Thr Glu Gly Thr Phe
      355          360          365
Pro Pro Phe Pro Ala Asp Ile Tyr Thr Ile Pro Gln Tyr Gly Tyr Cys
      370          375          380
Thr Leu Asn Tyr Asn Asn Glu Ala Val Asp Arg Ser Ala Phe Tyr Cys
      385          390          395          400
Leu Asp Tyr Phe Pro Ser Asp Met Leu Arg Thr Gly Asn Asn Phe Glu
      405          410          415

```

Phe Thr Tyr Thr Phe Glu Asp Val Pro Phe His Ser Met Phe Ala His  
 420 425 430  
 Asn Gln Thr Leu Asp Arg Leu Met Asn Pro Leu Val Asp Gln Tyr Leu  
 435 440 445  
 Trp Ala Phe Ser Ser Val Ser Gln Ala Gly Ser Ser Gly Arg Ala Leu  
 450 455 460  
 His Tyr Ser Arg Ala Thr Lys Thr Asn Met Ala Ala Gln Tyr Arg Asn  
 465 470 475 480  
 Trp Leu Pro Gly Pro Phe Phe Arg Asp Gln Gln Ile Phe Thr Gly Ala  
 485 490 495  
 Ser Asn Ile Thr Lys Asn Asn Val Phe Ser Val Trp Glu Lys Gly Lys  
 500 505 510  
 Gln Trp Glu Leu Asp Asn Arg Thr Asn Leu Met Gln Pro Gly Pro Ala  
 515 520 525  
 Ala Ala Thr Thr Phe Ser Gly Glu Pro Asp Arg Gln Ala Met Gln Asn  
 530 535 540  
 Thr Leu Ala Phe Ser Arg Thr Val Tyr Asp Gln Thr Thr Ala Thr Thr  
 545 550 555 560  
 Asp Arg Asn Gln Ile Leu Ile Thr Asn Glu Asp Glu Ile Arg Pro Thr  
 565 570 575  
 Asn Ser Val Gly Ile Asp Ala Trp Gly Ala Val Pro Thr Asn Asn Gln  
 580 585 590  
 Ser Ile Val Thr Pro Gly Thr Arg Ala Ala Val Asn Asn Gln Gly Ala  
 595 600 605  
 Leu Pro Gly Met Val Trp Gln Asn Arg Asp Ile Tyr Pro Thr Gly Thr  
 610 615 620  
 His Leu Ala Lys Ile Pro Asp Thr Asp Asn His Phe His Pro Ser Pro  
 625 630 635 640  
 Leu Ile Gly Arg Phe Gly Cys Lys His Pro Pro Pro Gln Ile Phe Ile  
 645 650 655  
 Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Glu Thr Phe Gln Thr Ala  
 660 665 670  
 Lys Val Ala Ser Phe Ile Asn Gln Tyr Ser Thr Gly Gln Cys Thr Val  
 675 680 685  
 Glu Ile Phe Trp Glu Leu Lys Lys Glu Thr Ser Lys Arg Trp Asn Pro  
 690 695 700  
 Glu Ile Gln Phe Thr Ser Asn Phe Gly Asn Ala Ala Asp Ile Gln Phe  
 705 710 715 720  
 Ala Val Ser Asp Thr Gly Ser Tyr Ser Glu Pro Arg Pro Ile Gly Thr  
 725 730 735  
 Arg Tyr Leu Thr Lys Pro Leu  
 740

&lt;210&gt; 12

&lt;211&gt; 1797

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
 synthetic construct

&lt;400&gt; 12

acggctccga	ccggagacaa	gcggaagggc	gaagacgaac	ctcgtttgcc	cgacacttct	60
tcacagactc	ccaagaaaaa	caagaagcct	cgcaaggaaa	gaccttccgg	cggggcagaa	120
gatccgggcg	aaggcacctc	ttccaacgct	ggagcagcag	caccgcctc	tagtgtggga	180
tcactatca	tggctgaagg	aggtggcggc	ccagtgggcg	atgcaggcca	gggtgccgat	240
ggagtgggca	attcctccgg	aaattggcat	tgcgattccc	aatggctgga	aaacggagtc	300
gtcactcgaa	ccaccgaac	ctgggtcttg	cccagctaca	acaaccacct	gtacaaacga	360
atccaaggac	ccagcggagg	cgacaacaac	aacaaattct	ttggattcag	caccccctgg	420
ggatactttg	actacaatcg	attccactgc	cacttttccc	cgcgagactg	gcaacgactc	480



```

atcaacaaca actggggcat ccgtcccaaa gcgatgcgct ttagactctt taacatccag      540
gttaaagagg tcacggtcca agacttcaac accaccatcg gcaacaacct caccagtacg      600
gtccaggctt ttgcggacaa ggactaccaa ctgccgtacg tcctcggatc ggctaccgaa      660
ggcaccttcc cgccgttccc agcggatata tacacgatcc cgcagtacgg gtactgcacg      720
ctaaactaca acaacgaggc ggtggatcgt tcggccttct actgtctgga ctactttccc      780
tcagacatgc tgcggacagg aaataacttt gagtttactt acaccttcga ggacgttcct      840
ttccatagca tgtttgcccc caaccagacg ctgaccgggc tgatgaatcc cctcgtggat      900
cagtacctct gggctttcag ctccgtcagc caagcaggct catctggacg agctcttcat      960
tactcgcggg cgactaaaac caacatggcg gctcaatata ggaactgggt acctgggcct     1020
ttcttcctgt atcagcaaat ctttacgggc gctagcaaca tcactaaaaa taacgtcttt     1080
agcgtttggg aaaaaggcaa gcaatgggaa ctcgacaatc ggaccaacct aatgcagccc     1140
ggtctcgcgg cagcgaccac ctttagcgga gaactgacc gtcaagccat gcaaaacacg     1200
ctggctttta acaggaccgt ctacgatcaa acgaccgcca cgaccgatcg taaccagata     1260
ctcatcacca acgaagacga aatcagaccc accaactcgg tcggtatcga cgcgtgggga     1320
gcagttccca ccaacaacca gtcgatcgtg acccccggca ctcgcgcggc cgtcaacaat     1380
caagggcgcg tttccgggat ggtgtggcaa aacagagaca ttacacctac agggacccat     1440
ttggccaaaa ttcccgacac tgacaatcac ttecatcgtg ccccgcttat tgggcggttt     1500
ggctgcaagc atccccctcc ccagattttc attaaaaaca caccgctccc tgccaacctt     1560
tcggaaacgt tccagacggc caaagtggcc tccttcacga accagtactc gaccggacag     1620
tgcaccgtcg aaatcttttg ggaactcaag aaggaaacct ccaagcgtcg gaaccccgaa     1680
atccagttca cctccaactt tggcaacgcg gccgacatcc agtttgccgt ctccgacacg     1740
ggatcctatt ccgaacctcg tcccatcggt acccgttacc ttaccaaacc tctgtaa      1797

```

&lt;210&gt; 13

&lt;211&gt; 598

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 13

```

Thr Ala Pro Thr Gly Asp Lys Arg Lys Gly Glu Asp Glu Pro Arg Leu
 1              5              10              15
Pro Asp Thr Ser Ser Gln Thr Pro Lys Lys Asn Lys Lys Pro Arg Lys
      20              25              30
Glu Arg Pro Ser Gly Gly Ala Glu Asp Pro Gly Glu Gly Thr Ser Ser
      35              40              45
Asn Ala Gly Ala Ala Ala Pro Ala Ser Ser Val Gly Ser Ser Ile Met
      50              55              60
Ala Glu Gly Gly Gly Gly Pro Val Gly Asp Ala Gly Gln Gly Ala Asp
      65              70              75              80
Gly Val Gly Asn Ser Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu
      85              90              95
Glu Asn Gly Val Val Thr Arg Thr Thr Arg Thr Trp Val Leu Pro Ser
      100             105             110
Tyr Asn Asn His Leu Tyr Lys Arg Ile Gln Gly Pro Ser Gly Gly Asp
      115             120             125
Asn Asn Asn Lys Phe Phe Gly Phe Ser Thr Pro Trp Gly Tyr Phe Asp
      130             135             140
Tyr Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu
      145             150             155             160
Ile Asn Asn Asn Trp Gly Ile Arg Pro Lys Ala Met Arg Phe Arg Leu
      165             170             175
Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Phe Asn Thr Thr
      180             185             190
Ile Gly Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ala Asp Lys Asp
      195             200             205
Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala Thr Glu Gly Thr Phe Pro
      210             215             220

```

Pro Phe Pro Ala Asp Ile Tyr Thr Ile Pro Gln Tyr Gly Tyr Cys Thr  
 225 230 235 240  
 Leu Asn Tyr Asn Asn Glu Ala Val Asp Arg Ser Ala Phe Tyr Cys Leu  
 245 250 255  
 Asp Tyr Phe Pro Ser Asp Met Leu Arg Thr Gly Asn Asn Phe Glu Phe  
 260 265 270  
 Thr Tyr Thr Phe Glu Asp Val Pro Phe His Ser Met Phe Ala His Asn  
 275 280 285  
 Gln Thr Leu Asp Arg Leu Met Asn Pro Leu Val Asp Gln Tyr Leu Trp  
 290 295 300  
 Ala Phe Ser Ser Val Ser Gln Ala Gly Ser Ser Gly Arg Ala Leu His  
 305 310 315 320  
 Tyr Ser Arg Ala Thr Lys Thr Asn Met Ala Ala Gln Tyr Arg Asn Trp  
 325 330 335  
 Leu Pro Gly Pro Phe Phe Arg Asp Gln Gln Ile Phe Thr Gly Ala Ser  
 340 345 350  
 Asn Ile Thr Lys Asn Asn Val Phe Ser Val Trp Glu Lys Gly Lys Gln  
 355 360 365  
 Trp Glu Leu Asp Asn Arg Thr Asn Leu Met Gln Pro Gly Pro Ala Ala  
 370 375 380  
 Ala Thr Thr Phe Ser Gly Glu Pro Asp Arg Gln Ala Met Gln Asn Thr  
 385 390 395 400  
 Leu Ala Phe Ser Arg Thr Val Tyr Asp Gln Thr Thr Ala Thr Thr Asp  
 405 410 415  
 Arg Asn Gln Ile Leu Ile Thr Asn Glu Asp Glu Ile Arg Pro Thr Asn  
 420 425 430  
 Ser Val Gly Ile Asp Ala Trp Gly Ala Val Pro Thr Asn Asn Gln Ser  
 435 440 445  
 Ile Val Thr Pro Gly Thr Arg Ala Ala Val Asn Asn Gln Gly Ala Leu  
 450 455 460  
 Pro Gly Met Val Trp Gln Asn Arg Asp Ile Tyr Pro Thr Gly Thr His  
 465 470 475 480  
 Leu Ala Lys Ile Pro Asp Thr Asp Asn His Phe His Pro Ser Pro Leu  
 485 490 495  
 Ile Gly Arg Phe Gly Cys Lys His Pro Pro Pro Gln Ile Phe Ile Lys  
 500 505 510  
 Asn Thr Pro Val Pro Ala Asn Pro Ser Glu Thr Phe Gln Thr Ala Lys  
 515 520 525  
 Val Ala Ser Phe Ile Asn Gln Tyr Ser Thr Gly Gln Cys Thr Val Glu  
 530 535 540  
 Ile Phe Trp Glu Leu Lys Lys Glu Thr Ser Lys Arg Trp Asn Pro Glu  
 545 550 555 560  
 Ile Gln Phe Thr Ser Asn Phe Gly Asn Ala Ala Asp Ile Gln Phe Ala  
 565 570 575  
 Val Ser Asp Thr Gly Ser Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg  
 580 585 590  
 Tyr Leu Thr Lys Pro Leu  
 595

&lt;210&gt; 14

&lt;211&gt; 1608

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
 synthetic construct

&lt;400&gt; 14

atggctgaag gaggtggcgg cccagtgggc gatgcaggcc aggggtgccga tggagtgggc 60  
 aattctccg gaaattggca ttgcgattcc caatggctgg aaaacggagt cgtcactcga 120

```

accacccgaa cctgggtctt gccagctac aacaaccacc tgtacaaacg aatccaagga      180
cccagcggag gcgacaacaa caacaaattc tttggattca gcaccccttg gggatacttt      240
gactacaatc gattccactg ccacttttcc ccgcgagact ggcaacgact catcaacaac      300
aactggggca tccgtcccaa agcgatgcgc tttagactct ttaacatcca ggttaaagag      360
gtcacgggtc aagacttcaa caccaccatc ggcaacaacc tcaccagtac ggtccaggtc      420
tttgcggaac aggactacca actgccgtac gtccctcggt cggtaccga aggcaccttc      480
ccgcctgtcc cagcggatat ctacacgac ccgcagtagc ggtactgcac gctaaactac      540
aacaacgagg cgggtggatcg ttcggccttc tactgtctgg actactttcc ctcagacatg      600
ctgcggaacg gaaataactt tgagtttact tacaccttcg aggacgttcc tttccatagc      660
atgtttgccc acaaccagac gctagaccgg ctgatgaatc ccctcgtgga tcagtacctc      720
tgggctttca gctccgtcag ccaagcaggc tcactctggac gagctcttca ttactcgagg      780
gcgactaaaa ccaacatggc ggctcaatat aggaactggg tacctgggac tttcttccgt      840
gatcagcaaa tctttacggg cgctagcaac atcactaaaa ataacgtctt tagcgtttgg      900
gaaaaaggca agcaatggga actcgacaat cggaccaacc taatgcagcc cggctcctgcg      960
gcagcgacca cttttagcgg agaactgac cgtcaagcca tgcaaaacac gctggctttt     1020
agcaggaccg tctacgatca aacgaccgcc acgaccgac gtaaccagat actcatcacc     1080
aacgaagacg aatcagacc caccaactcg gtcggtatcg acgctggggg agcagttccc     1140
accaacaacc agtcgatcgt gaccccgggc actcgcgcgg ccgtcaacaa tcaagggggc     1200
cttcccggga tgggtgggca aaacagagac atttacccta cagggacca tttggccaaa     1260
attcccgaac ctgacaatca cttccatcgg tccccgctta ttgggcggtt tggctgcaag     1320
catccccctc cccagatttt cattaaaaac acaccgctcc ctgccaaacc ttcggaaacg     1380
ttccagacgg ccaaagtggc ctcttcacac aaccagtact cgaccggaca gtgcaccgtc     1440
gaaatctttt gggaaactcaa gaaggaaacc tccaagcgct ggaaccccgga aatccagttc     1500
acctccaact ttggcaacgc ggccgacatc cagtttgccg tctccgacac gggatcctat     1560
tccgaacctc gtcccatcgg taccggttac cttaccaaac ctctgtaa      1608

```

&lt;210&gt; 15

&lt;211&gt; 535

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
synthetic construct

&lt;400&gt; 15

```

Met Ala Glu Gly Gly Gly Gly Pro Val Gly Asp Ala Gly Gln Gly Ala
 1           5           10           15
Asp Gly Val Gly Asn Ser Ser Gly Asn Trp His Cys Asp Ser Gln Trp
          20           25           30
Leu Glu Asn Gly Val Val Thr Arg Thr Thr Arg Thr Trp Val Leu Pro
          35           40           45
Ser Tyr Asn Asn His Leu Tyr Lys Arg Ile Gln Gly Pro Ser Gly Gly
          50           55           60
Asp Asn Asn Asn Lys Phe Phe Gly Phe Ser Thr Pro Trp Gly Tyr Phe
65           70           75           80
Asp Tyr Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg
          85           90           95
Leu Ile Asn Asn Asn Trp Gly Ile Arg Pro Lys Ala Met Arg Phe Arg
          100          105          110
Leu Phe Asn Ile Gln Val Lys Glu Val Thr Val Gln Asp Phe Asn Thr
          115          120          125
Thr Ile Gly Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ala Asp Lys
          130          135          140
Asp Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala Thr Glu Gly Thr Phe
145          150          155          160
Pro Pro Phe Pro Ala Asp Ile Tyr Thr Ile Pro Gln Tyr Gly Tyr Cys
          165          170          175
Thr Leu Asn Tyr Asn Asn Glu Ala Val Asp Arg Ser Ala Phe Tyr Cys
          180          185          190

```

Leu Asp Tyr Phe Pro Ser Asp Met Leu Arg Thr Gly Asn Asn Phe Glu  
 195 200 205  
 Phe Thr Tyr Thr Phe Glu Asp Val Pro Phe His Ser Met Phe Ala His  
 210 215 220  
 Asn Gln Thr Leu Asp Arg Leu Met Asn Pro Leu Val Asp Gln Tyr Leu  
 225 230 235 240  
 Trp Ala Phe Ser Ser Val Ser Gln Ala Gly Ser Ser Gly Arg Ala Leu  
 245 250 255  
 His Tyr Ser Arg Ala Thr Lys Thr Asn Met Ala Ala Gln Tyr Arg Asn  
 260 265 270  
 Trp Leu Pro Gly Pro Phe Phe Arg Asp Gln Gln Ile Phe Thr Gly Ala  
 275 280 285  
 Ser Asn Ile Thr Lys Asn Asn Val Phe Ser Val Trp Glu Lys Gly Lys  
 290 295 300  
 Gln Trp Glu Leu Asp Asn Arg Thr Asn Leu Met Gln Pro Gly Pro Ala  
 305 310 315 320  
 Ala Ala Thr Thr Phe Ser Gly Glu Pro Asp Arg Gln Ala Met Gln Asn  
 325 330 335  
 Thr Leu Ala Phe Ser Arg Thr Val Tyr Asp Gln Thr Thr Ala Thr Thr  
 340 345 350  
 Asp Arg Asn Gln Ile Leu Ile Thr Asn Glu Asp Glu Ile Arg Pro Thr  
 355 360 365  
 Asn Ser Val Gly Ile Asp Ala Trp Gly Ala Val Pro Thr Asn Asn Gln  
 370 375 380  
 Ser Ile Val Thr Pro Gly Thr Arg Ala Ala Val Asn Asn Gln Gly Ala  
 385 390 395 400  
 Leu Pro Gly Met Val Trp Gln Asn Arg Asp Ile Tyr Pro Thr Gly Thr  
 405 410 415  
 His Leu Ala Lys Ile Pro Asp Thr Asp Asn His Phe His Pro Ser Pro  
 420 425 430  
 Leu Ile Gly Arg Phe Gly Cys Lys His Pro Pro Pro Gln Ile Phe Ile  
 435 440 445  
 Lys Asn Thr Pro Val Pro Ala Asn Pro Ser Glu Thr Phe Gln Thr Ala  
 450 455 460  
 Lys Val Ala Ser Phe Ile Asn Gln Tyr Ser Thr Gly Gln Cys Thr Val  
 465 470 475 480  
 Glu Ile Phe Trp Glu Leu Lys Lys Glu Thr Ser Lys Arg Trp Asn Pro  
 485 490 495  
 Glu Ile Gln Phe Thr Ser Asn Phe Gly Asn Ala Ala Asp Ile Gln Phe  
 500 505 510  
 Ala Val Ser Asp Thr Gly Ser Tyr Ser Glu Pro Arg Pro Ile Gly Thr  
 515 520 525  
 Arg Tyr Leu Thr Lys Pro Leu  
 530 535

&lt;210&gt; 16

&lt;211&gt; 120

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence; note =  
 synthetic construct

&lt;400&gt; 16

tggccagttt ccaagacagg ctgcctcgct cactcgggcc ggggccccaa aggggcccct 60  
 agcgaccgct tcgcggtcgc ggcccagtg agcgagcgag cctgtcttgg aaactggcca 120

<210> 17  
<211> 120  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 17  
accggtcaaa ggttctgtcc gagcgagcga gtgagcccg ccccggcgtt tccccgggga 60  
tcgctgggga agcgccagcg ccgggctcac tcgctcgctc ggacagaacc ttgaccggt 120

<210> 18  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 18  
gcactccggt gagtaatgc cg 22

<210> 19  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 19  
cggcattacc tcaccggagt gc 22

<210> 20  
<211> 7  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 20  
ccggtcg 7

<210> 21  
<211> 16  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 21  
cgagtgagcg agcgag 16

<210> 22  
<211> 101  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 22  
tcacgtggtc gggaatggga acgggaaatc tcgcgagaac gtaaacaat ataagacggc 60  
gccacacggc gctgcgtcat acgcgcgcgc gcaccggcga g 101

<210> 23  
<211> 252  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 23  
aatctcgggg gcgcgaacaa ggtccggggc gagtcgtata ttcccgccta cctgatcccg 60  
aaacagcaac cggaagtgc gtgggctgg actaacgtgc ccgagtatat aaaagcgtgc 120  
ttgcaccgag aactgcgtgc cagtctcgcg cgacttcact tcgaggaggc gggcgtctcg 180  
caatccaagg aaaatctcgc gagaactgca gacggcgctc ccgtgatgcc gaccgcgtc 240  
agcaaacgct ac 252

<210> 24  
<211> 196  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence; note =  
synthetic construct

<400> 24  
gtaaggtgac caaacaggag gtgcgtcaat tcttcagggtg gtctcaggat cacctgaccc 60  
ctgtgatccc agaattccta gtgcggaagg cggagtctcg caaaagaccc gcccttccg 120  
gggaaggcta tataagcccg acaaagcggc ccgcgctcgc agagcagcag caggcgtcgg 180  
agagcgcgga cccggt 196